# Synthetic Studies of Therapeutic Targets and Their Application to the Development of Mirror-Image Single-Domain Antibodies

2023

Keisuke Aoki

# Contents

# Preface

Biological systems predominantly use one enantiomer, which is termed biological homochirality.[1] In nature, DNA and RNA sugars are the D-configuration, whereas the proteinogenic amino acids are the L-configuration (apart from achiral glycine) (Figure 1). In addition to classical small molecule drugs, naturally existing biomolecules, such as proteins and nucleic acids, are used as therapeutics.[2,3] In particular, proteins serve as an important category of therapeutic agents because of their high bioactivity and specificity.[2] The first commercially available recombinant protein was human insulin, which remains the major therapy for treating diabetes mellitus type I and type II.[4] Other protein therapeutics, including erythropoietin, granulocyte colony-stimulating factor (G-CSF), and interferons, are used for hematopoiesis and immune regulation.[5–7] These therapeutic proteins augment the function of the corresponding endogenous proteins.

Among protein therapeutics, antibodies have been approved widely for treating cancer, autoimmune diseases, and viral infections.[8] For example, the anti-human epidermal growth factor receptor 2 (HER2) monoclonal antibody (mAb), trastuzumab, and the anti-vascular endothelial growth factor A (VEGF-A) mAb, bevacizumab, exhibit anti-tumor activity by inhibiting ligand-receptor interactions vital for tumor survival and propagation.[9] The anti-tumor necrosis factor (TNF) mAb, adalimumab, and the anti-interleukin-6 receptor (IL-6R) mAb, tocilizumab, are essential in treating rheumatoid arthritis.[10] Recently, several mAbs targeting the spike protein of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) were rapidly developed and approved to fight the coronavirus disease 2019 (COVID-19) pandemic.[11]

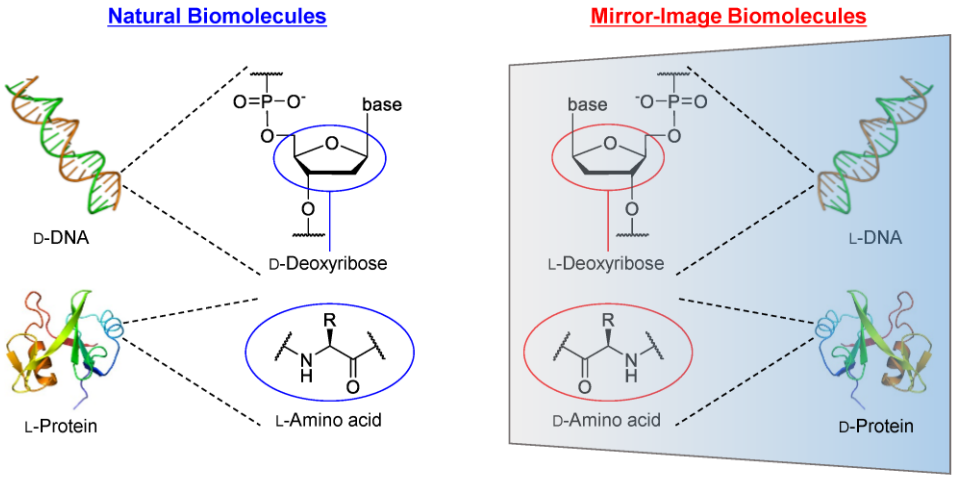


**Figure 1.** Representative biomolecules and their mirror-image isomers.

Although the majority of approved antibodies are in the IgG format, several alternatives are emerging, including antibody-drug conjugates (ADCs), bispecific antibodies, and fragment antibodies (Figure 2).[8] ADCs are complex molecules with the targeting ability of antibodies connected to a cytotoxic payload via a linker. Bispecific antibodies simultaneously bind two different antigens. For example, the bispecific antibody emicizumab, which is approved for treating hemophilia A, recognizes both the enzyme factor IXa and the substrate factor X to promote colocalization of the enzyme-substrate complex.[12] Engineered antibody fragments, such as antigen-binding fragments (Fab), single-chain variable fragments (scFv), and variable domains of the heavy chain antibody (VHH) are also promising small protein scaffolds for pharmaceutical applications that may reproduce the high affinity and specificity of conventional antibodies.[13] One advantage of these fragments is that they penetrate tissues and tumors more rapidly and deeply than mAbs because of their small size.[14] The design flexibility for various bioconjugates and low production costs are also strengths of these fragments.[15]

However, the clinical efficacy of protein therapeutics often declines during prolonged administration.[16,17] For example, during the treatment of hemophilia A patients with exogenous factor VIII protein, the emergence of biomolecular inhibitors was reported in approximately 30% of patients, which diminished the therapeutic effect of the protein therapeutic.[16] In a cohort study of anti-TNF mAb administration, loss of response was observed in 25% of patients.[17] In addition, these therapies are associated with adverse effects, such as infusion reactions, anaphylaxis and pure red cell aplasia.[18] Both the reduction in efficacy and incidence of adverse events are often linked to the generation of anti-drug antibodies (ADAs) caused by immunogenic responses (Figure 3).[19] Neutralizing ADAs (NAbs) block drug binding to the pharmacologically active site, whereas non-neutralizing ADAs bind the drug molecules and indirectly reduce drug levels by increasing drug clearance via immune complex formation.[20] In some cases, ADAs may cause immune complex-mediated diseases, such as serum sickness and thromboembolic events.[21,22] Therefore, it is crucial to reduce the immunogenicity of protein therapeutics.

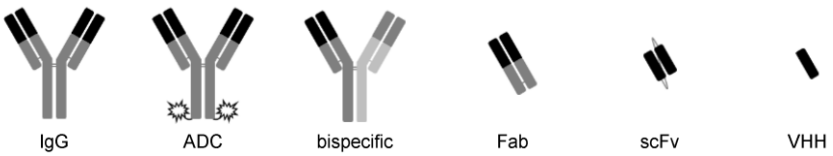


**Figure 2.** Representative molecular formats of antibody-based therapeutics.

Using native human proteins or the humanization of antibodies is a robust approach to reducing the possible risk of ADA production.[23] Human insulin was less immunogenic than the previously used animal insulin isolated from the bovine or porcine pancreas.[24] Similarly, several technologies have been developed to address the safety issues of murine monoclonal antibodies. Technologies for generating humanized or human monoclonal antibodies, including complementarity-determining region (CDR) grafting, phage display, and transgenic mouse platforms may contribute to reducing the immunogenicity of the mouse-derived protein sequences.[25–27] Eliminating predicted T cell epitopes is another promising strategy to minimize the immune response.[28] During the development of the bispecific engineered antibody emicizumab, the antibody sequence was optimized to reduce the immunogenicity risk using an *in silico* T cell epitope prediction system and improve the activity and physicochemical properties.[29]

Mirror-image biomolecules (e.g., D-peptides and D-proteins) are promising biomolecules to circumvent the potential risk of immunogenicity and rapid degradation.[30] Because D-peptides and D-proteins consisting of D-amino acids are unlikely to be degradable substrates for proteolytic enzymes in antigen-presenting cells (APCs), ADA generation via T cell activation caused by major histocompatibility complex (MHC) presentation can be avoided (Figure 3).[31] In 1993, Berg *et al.* demonstrated the first direct evidence that a folded D-protein has low immunogenicity and extended half-life *in vivo*.[31] In this study, both enantiomers of rubredoxin (45 a.a.) were synthesized and administered with alum adjuvant in mice. In their comparative analysis, the D-protein did not induce a specific IgG antibody response, whereas the L-protein induced a robust immune response.

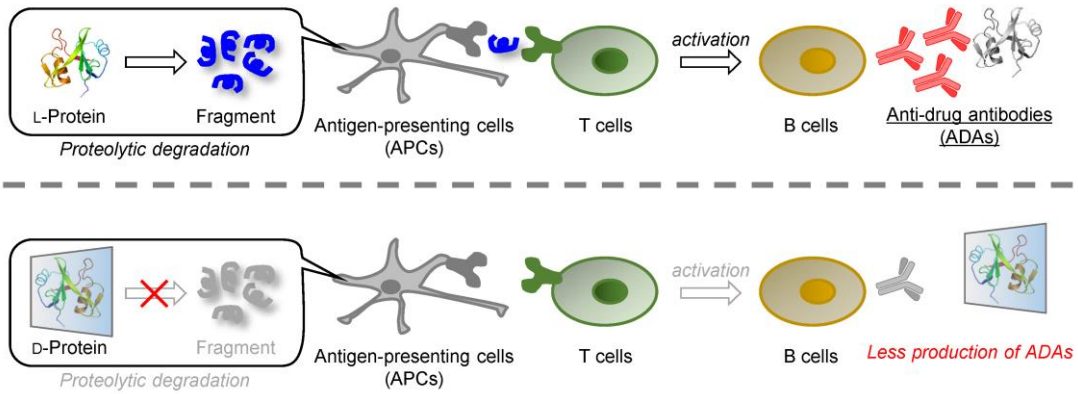


**Figure 3.** Schematic representations of immune responses that occur during administration of both enantiomers of protein therapeutics. Mirror-image proteins (D-proteins) are expected to be less susceptive to proteolytic degradation in antigen-presenting cells (APCs), thus resulting in lower production of anti-drug antibodies (ADAs).

Classically, target-binding D-peptides were designed by retro-inversion, which involves the chirality of each residue and the order of the amino acid sequence being inverted from the parent L-peptide.[32] Although this approach succeeded in some unstructured peptides,[33,34] the binding activities of retro-inverso peptides with α-helix secondary structures usually decreased because of the differences in topology between L- and D-peptides.[35]

An advanced method to discover mirror-image molecules with desired pharmacological effects was reported in 1996 by Kim and coworkers.[36] They developed the mirror-image phage display (MIPD) technology and identified D-peptides binding to Src homology 3 domain of c-Src kinase (Src-SH3) (Figure 4). In this approach, the synthetic D-enantiomer of the target protein(s) (D-target) was used to screen against the L-peptide-displaying phage library with high diversity. Based on the hit sequences of L-peptides binding to the D-target identified from this "mirror-image screening", they synthesized the mirror-image peptides (D-peptides) of the hit sequences. The resulting D-peptides bound to the L-target with equipotent affinity, revealing that MIPD is an alternative to the standard screening protocol using native target proteins (L-targets) for the unavailable D-peptide library. Since the first report of MIPD, this approach has been adopted to explore novel D-peptides for a range of therapeutic targets.[37–47] Notably, the effects of several D-peptide therapeutic candidates, including inhibitors of viral entry and amyloid β (Aβ) fibril formation, have been evaluated in clinical trials.[48,49]

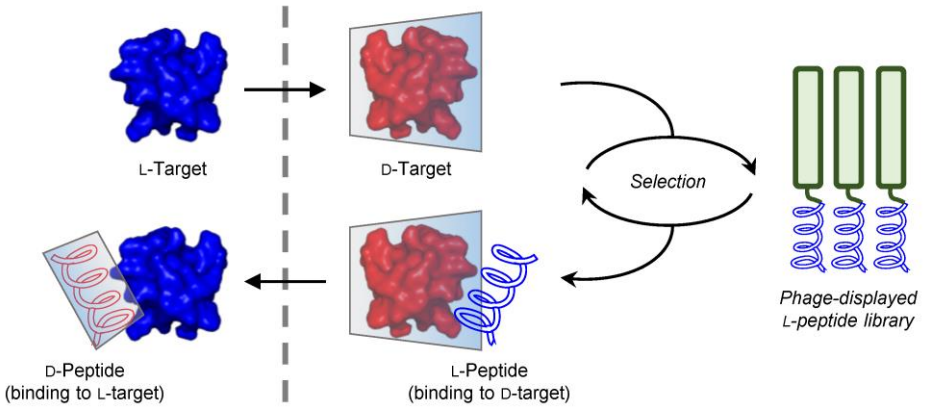


**Figure 4.** Schematic representation of mirror-image phage display (MIPD) to identify D-peptide ligand(s) that bind to a therapeutic target (L-target). In the first step, a mirror-image isomer of the target protein (D-target) is synthesized. In the second step, the D-target is screened against a phage-displayed L-peptide library and L-peptide(s) binding to the D-target is selected. In the third step, the mirror-image isomer of the selected L-peptide(s) is synthesized using D-amino acids, which will bind to the L-target.

The scope of mirror-image screening is not limited to only identifying D-peptides (Figure 5). Kent and Sidhu *et al.* developed D-proteins specific for VEGF-A.[50–52] This development was achieved using libraries of small proteins containing some random sequences, including the B1 domain of streptococcal protein G (56 a.a.), the GA domain of protein G (53 a.a.), and the Z domain of protein A (58 a.a.). The identified D-protein blocked VEGF-A-induced vascular leakages, inhibited tumor growth *in vivo*, and did not elicit ADA production.[52] Mirror-image RNA aptamers (L-RNA aptamers), termed Spiegelmers, were also developed to address the disadvantages of standard RNA aptamers (D-RNA aptamers), such as low *in vivo* stability because of nuclease digestion and immunogenicity.[53] The potential of Spiegelmers as therapeutics has been evaluated in several clinical trials. The author's group applied a mirror-image screening system to discover drug candidates from the unexplored mirror-image structures of natural products by combining this system with a chemical array technology.[54] Using this approach, a novel inhibitor against mouse double minute 2 homolog (MDM2)-p53 from a mirror-image library of natural products was identified.[54]

To expand the scope of the mirror-image screening technology, the author focused on VHH in camelids, which is a single-domain antibody with favorable properties, including high stability and improved solubility.[55] The author hypothesized that a mirror-image form of VHH (D-VHH) overcomes the reported ADA issue of conventional protein therapeutics to provide a novel D-protein-based scaffold with high-affinity target binding

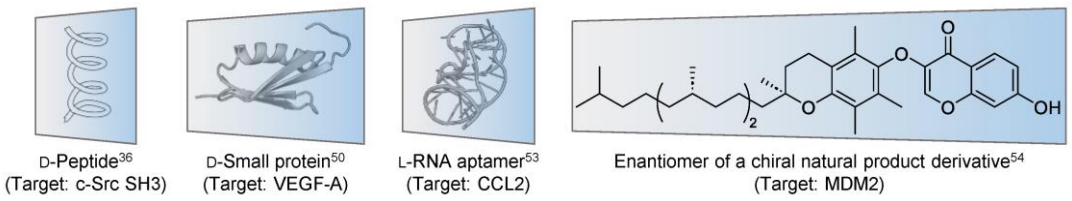


**Figure 5.** Representative mirror-image molecules identified in previous studies using mirror-image screening technology.

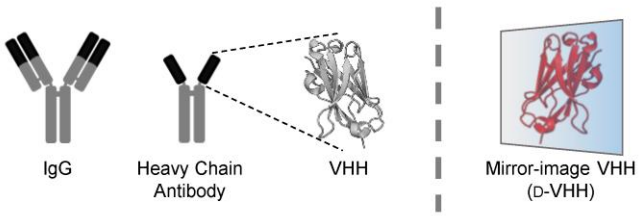


**Figure 6.** Schematic diagram of the full-length antibody (IgG format), heavy chain antibody in *Camelidae*, VHH, and mirror-image VHH (D-VHH). VHH is the variable domain of the heavy chain antibody in *Camelidae*.

and low immunogenicity (Figure 6). Thus, the author planned the development of the screening platform to identify novel therapeutics from the virtual D-VHH library in a mirror.

A bottleneck for the broad application of mirror-image screening is the preparation of D-target proteins by chemical synthesis.[30] Because D-proteins cannot be prepared using conventional recombinant technology, several chemical methods have been developed to prepare mirror-image biomolecules.[56,57] Solid-phase peptide synthesis (SPPS) is a common approach for synthesizing short peptides; however, the accessible size of the peptide segments is limited (typically ~50 amino acids).[56] The advancement of ligation techniques, such as native chemical ligation (NCL), has facilitated the synthesis of larger proteins.[57] In chemical protein synthesis, optimizing a suitable synthetic route for constructing the full-length sequence and investigating the refolding conditions to reproduce the biological functions are generally required.[58] In some cases, installation and removal of solubilizing tags are required to address the poor solubility of some hydrophobic peptides.[59] Chemical protein synthesis often becomes more challenging as the number of amino acid residues increases.[58,59] Thus, examples of target proteins applied to mirror-image screening have been limited to only small proteins, such as Aβ (42 a.a.), MDM2 (85 a.a.), and VEGF-A (108 a.a.).[36–47,50–54] To expand the application of mirror-image screening technology toward unexplored medium-sized target proteins (> 150 a.a.), the author investigated the synthesis of several therapeutic targets.

In this thesis, the author describes the development of the screening process from a mirror-image library of single-domain antibodies. The author also investigates the chemical synthesis of therapeutic targets for applications in mirror-image screening.

In Chapter 1, Section 1, the author describes the chemical synthesis of mirror-image VHH proteins (~120 a.a.) and investigates their characteristics, including biodistribution and immunogenicity.

In Chapter 1, Section 2, the author describes the construction of phage-displayed VHH libraries and their application to exploring the mirror-image VHH specific for VEGF-A.

In Chapter 2, Section 1, the author describes the chemical synthesis of the full-length core protein of the hepatitis B virus (HBV) (183 a.a.), which is an attractive target for treating chronic HBV infections.

In Chapter 2, Section 2, the author describes the chemical synthesis of interleukin-6 (183 a.a.), a target protein for immunological and inflammatory diseases, and its application to mirror-image screening.

References


1. Blackmond, D. G. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032540.
2. Leader, B.; Baca, Q. J.; Golan, D. E. *Nat. Rev. Drug Discov.* **2008**, *7*, 21–39.
3. Kulkarni, J. A.; Witzigmann, D.; Thomson, S. B.; Chen, S.; Leavitt, B. R.; Cullis, P. R.; van der Meel, R. *Nat. Nanotechnol.* **2021**, *16*, 630–643.
4. Mathieu, C.; Martens, P.-J.; Vangoitsenhoven, R. *Nat. Rev. Endocrinol.* **2021**, *17*, 715–725.
5. Erslev, A. J. *N. Engl. J. Med.* **1991**, *324*, 1339–1344.
6. Lieshke, G. J.; Burgess, A. W. *N. Engl. J. Med.* **1992**, *327*, 99–106.
7. Teijaro, J. R. *Curr. Opin Virol.* **2016**, *16*, 31–40.
8. Carter, P. J.; Rajpal, A. *Cell* **2022**, *185*, 2789–2805.
9. Zinn, S.; Vazquez-Lombardi, R.; Zimmermann, C.; Sapra, P.; Jermutus, L.; Christ, D. *Nat. Cancer* **2023**, *4*, 165–180.
10. Aletaha, D. *JAMA* **2018**, *320*, 1360–1372.
11. Corti, D.; Purcell, L. A.; Snell, G.; Veesler, D. *Cell* **2021**, *184*, 3086–3108.
12. Lenting, P. J.; Denis, C. V.; Christophe, O. D. *Blood* **2017**, *130*, 2463–2468.
13. Holliger, P.; Hudson, P. J. *Nat. Biotechnol.* **2005**, *23*, 1126–1136.
14. Debie, P.; Lafont, C.; Defrise, M.; Hansen, I.; van Willigen, D. M.; van Leeuwen, F. W. B.; Gijsbers, R.; D'Huyvetter, M.; Devoogdt, N.; Lahoutte, T.; Mollard, P.; Hernot, S. *J. Controlled Release* **2020**, *317*, 34–42.
15. Steeland, S.; Vandenbroucke, R. E.; Libert, C. *Drug Discov. Today* **2016**, *21*, 1076–1113.
16. Wight, J.; Paisley, S. *Haemophilia* **2003**, *9*, 418–435.
17. Schultheiss, J. P. D.; Mahmoud, R.; Louwers, J. M.; van der Kaaij, M. T.; van Hellemondt, B. P.; van Boeckel, P. G.; Mahmmod, N.; Jharap, B.; Fidder, H. H.; Oldenburg, B. *Aliment. Pharmacol. Ther.* **2021**, *54*, 1298–1308.
18. Giezen, T. J.; Mantel-Teeuwisse, A. K.; Straus, S. M. J. M.; Schellekens, H.; Leufkens, H. G. M.; Egberts, A. C. G. *JAMA* **2008**, *300*, 1887–1896.
19. Dingman, R.; Balu-Iyer, S. V. *J. Pharm. Sci.* **2019**, *108*, 1637–1654.
20. Shankar, G.; Arkin, S.; Cocea, L.; Devanarayan, V.; Kirshner, S.; Kromminga, A.; Quarmby, V.; Richards, S.; Schneider, C. K.; Subramanyam, M.; Swanson, S.; Verthelyi, D.; Yim, S. *AAPS J.* **2014**, *16*, 658–673.
21. D'Arcy, C. A.; Mannik, M. *Arthritis. Rheum.* **2001**, *44*, 1717–1718.
22. Korswagen, L. A.; Bartelds, G. M.; Krieckaert, C. L. M.; Turkstra, F.; Nurmohamed, M. T.; van Schaardenburg, D.; Wijbrandts, C. A.; Tak, P. P.; Lems, W. F.; Dijkmans,

B. A. C.; van Vugt, R. M.; Wolbink, G. J. *Arthritis. Rheum.* **2011**, *63*, 877–883.
23. Hwang, W. Y. K.; Foote, J. *Methods* **2005**, *36*, 3–10.
24. Fineberg, S. E.; Galloway, J. A.; Fineberg, N. S.; Rathbun, M. J.; Hufferd, S. *Diabetologia* **1983**, *25*, 465–469.
25. Riechmann, L.; Clark, M.; Waldmann, H.; Winter, G. *Nature* **1988**, *332*, 323–327.
26. McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J. *Nature* **1990**, *348*, 552–554.
27. Lonberg, N.; Taylor, L. D.; Harding, F. A.; Trounstine, M.; Higgins, K. M.; Schramm, S. R.; Kuo, C.-C.; Mashayekh, R.; Wymore, K.; McCabe, J. G.; Munoz-O'Regan, D.; O'Donnell, S. L.; Lapachet, E. S. G.; Bengoechea, T.; Fishwild, D. M.; Carmack, C. E.; Kay, R. M.; Huszar, D. *Nature* **1994**, 368, 856–859.
28. King, C.; Garza, E.; Mazor, R.; Linehan, J. L.; Pastan, I.; Pepper, M.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 8577–8582.
29. Sampei, Z.; Igawa, T.; Soeda, T.; Okuyama-Nishida, Y.; Moriyama, C.; Wakabayashi, T.; Tanaka, E.; Muto, A.; Kojima, T.; Kitazawa, T.; Yoshihashi, K.; Harada, A.; Funaki, M.; Haraya, K.; Tachibana, T.; Suzuki, S.; Esaki, K.; Nabuchi, Y.; Hattori, K. *PLoS One* **2013**, *8*, e57479.
30. Harrison, K.; Mackay, A. S.; Kambanis, L.; Maxwell, J. W. C.; Payne, R. J. *Nat. Rev. Chem.* **2023**, *7*, 383–404.
31. Dintzis, H. M.; Symer, D. E.; Dintzis, R. Z.; Zawadzke, L. E.; Berg, J. M. *Proteins: Struct., Funct., Bioinf.* **1993**, *16*, 306–308.
32. Goodman, M.; Chorev, M. *Acc. Chem. Res.* **1979**, *12*, 1–7.
33. Ben-Yedidia, T.; Beignon, A.-S.; Partidos, C. D.; Muller, S.; Arnon, R. *Mol. Immunol.* **2002**, *39*, 323–331.
34. Li, H.; Kem, D. C.; Zhang, L.; Huang, B.; Liles, C.; Benbrook, A.; Gali, H.; Veitla, V.; Scherlag, B. J.; Cunningham, M. W.; Yu, X. *Hypertension* **2015**, *65*, 793–799.
35. Li C.; Pazgier, M.; Li, J.; Li, C.; Liu, M.; Zou, G.; Li, Z.; Chen, J.; Tarasov, S. G.; Lu, W.-Y.; Lu, W. *J. Biol. Chem.* **2010**, *285*, 19572–19581.
36. Schumacher, T. N. M.; Mayr, L. M.; Minor, D. L. Jr.; Milhollen, M. A.; Burgess, M. W.; Kim, P. S. *Science* **1996**, *271*, 1854–1857.
37. Eckert, D. M.; Malashkevich, V. N. Hong, L. H.; Carr, P. A.; Kim, P. S. *Cell* **1999**, *99*, 103–115.
38. Wiesehan, K.; Buder, K.; Linke, R. P.; Patt, S.; Stoldt, M.; Unger, E.; Schmitt, B.; Bucci, E.; Willbold, D. *ChemBioChem* **2003**, *4*, 748–753.
39. Liu, M.; Pazgier, M.; Li, C.; Yuan, W.; Li, C.; Lu, W. *Angew. Chem., Int. Ed.* **2010**, *49*, 3649–3652.

40. Chang, H.-N.; Liu, B.-Y.; Qi, Y.-K.; Zhou, Y.; Chen, Y.-P.; Pan, K.-M.; Li, W.-W.; Zhou, X.-M.; Ma, W.-W.; Fu, C.-Y.; Qi, Y.-M.; Liu, L.; Gao, Y.-F. *Angew. Chem., Int. Ed.* **2015**, *54*, 11760–11764.
41. Dammers, D.; Yolcu, D.; Kukuk, L.; Willbold, D.; Pickhardt, M.; Mandelkow, E.; Hom, A. H. C.; Sticht, H.; Malhis, M. N.; Will, N.; Schuster, J.; Funke, S. A. *PLoS One* **2016**, *11*, e0167432.
42. Li, Z.; Xie, J.; Peng, S.; Liu, A.; Wang, Y.; Lu, W.; Shen, J.; Li, C. *Bioconjug. Chem.* **2017**, *28*, 2167–2179.
43. Díaz-Perlas, C.; Varese, M.; Guardiola, S.; Sánchez-Navarro, M.; García, J.; Teixidó, M.; Giralt, E. *ChemBioChem* **2019**, *20*, 2079–2084.
44. Zhou, X.; Zuo, C.; Li, W.; Shi, W.; Zhou, X.; Wang, H.; Chen, S.; Du, J.; Chen, G.; Zhai, W.; Zhao, W.; Wu, Y.; Qi, Y.; Liu, L.; Gao, Y. *Angew. Chem., Int. Ed.* **2020**, *59*, 15114–15118.
45. Nonaka, M.; Mabashi-Asazuma, H.; Jarvis, D. L.; Yamasaki, K.; Akama, T. O.; Nagaoka, M.; Sasai, T.; Kimura-Takagi, I.; Suwa, Y.; Yaegashi, T.; Huang, C.-T.; Nishizawa-Harada, C.; Fukuda, M. N. *PLoS One* **2021**, *16*, e0241157.
46. Kokkwitz, P. E.; Mohrlüder, J.; Willbold, D. *Biomolecules* **2022**, *12*, 157.
47. Callahan, A. J.; Gandhesiri, S.; Travaline, T. L.; Salazar, L. L.; Hanna, S.; Lee, Y.-C.; Li, K.; Tokareva, O. S.; Swiecicki, J.-M.; Loas, A.; Verdine, G. L.; McGee, J. H.; Pentelute, B. L. *ChemRxiv* **2023**, https://doi.org/10.26434/chemrxiv-2023-x86xp.
48. Redman, J. S.; Francis, J. N.; Marquardt, R.; Papac, D.; Mueller, A. L.; Eckert, D. M.; Welch, B. D.; Kay, M. S. *Mol. Pharm.* **2018**, *15*, 1169–1179.
49. Kutzsche, J.; Jürgens, D.; Willuweit, A.; Adermann, K.; Fuchs, C.; Simons, S.; Windisch, M.; Hümpel, M.; Rossberg, W.; Wolzt, M.; Willbold, D. *Alzheimer's Dement.* **2020**, *6*, e12001.
50. Mandal, K.; Uppalapati, M.; Ault-Riché, D.; Kenney, J.; Lowitz, J.; Sidhu, S. S.; Kent, S. B. H. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 14779–14784.
51. Uppalapati, M.; Lee, D. J.; Mandal, K.; Li, H.; Miranda, L. P.; Lowitz, J.; Kenney, J.; Adams, J. J.; Ault-Riché, D.; Kent, S. B. H.; Sidhu, S. S. *ACS Chem. Biol.* **2016**, *11*, 1058–1065.
52. Marinec, P. S.; Landgraf, K. E.; Uppalapati, M.; Chen, G.; Xie, D.; Jiang, Q.; Zhao, Y.; Petriello, A.; Deshayes, K.; Kent, S. B. H.; Ault-Riche, D.; Sidhu, S. S. *ACS Chem. Biol.* **2021**, *16*, 548–556.
53. Vater, A.; Klussmann, S. *Drug Discov. Today* **2015**, *20*, 147–155 and the references cited therein.

54. Noguchi, T.; Oishi, S.; Honda, K.; Kondoh, Y.; Saito, T.; Ohno, H.; Osada. H.; Fujii, N. *Chem. Commun.* **2016**, *52*, 7653–7656.
55. Muyldermans, S. *Annu. Rev. Biochem.* **2013**, *82*, 775–797.
56. Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.
57. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.
58. Sun, H.; Brik, A. *Acc. Chem. Res.* **2019**, *52*, 3361–3371.
59. Mueller, L. K.; Baumruck, A. C.; Zhdanova, H.; Tietze, A. A. *Front. Bioeng. Biotechnol.* **2020**, *8*, 162.

Chapter 1. Development and Application of a Screening Platform for Mirror-Image Single Domain Antibodies

Section 1. Chemical Synthesis of Mirror-Image VHHs and Evaluation of Their Immunogenicity

*Summary: The author investigated the preparation protocols of mirror-image proteins of the variable domain of the heavy chain of the heavy chain antibody (D-VHHs) and their biological properties, including stereoselective target binding and immunogenicity. Initially, the author established a facile synthetic process of two model VHHs [anti-GFP VHH and PMP12A2h1 (monomeric VHH of caplacizumab)] and their mirror-image proteins by three-step native chemical ligations (NCLs) from four peptide segments. The folded synthetic VHHs (L-anti-GFP VHH and L-PMP12A2h1) bound to the target proteins (EGFP and vWF-A1 domain, respectively), while their mirror-image proteins (D-anti-GFP VHH and D-PMP12A2h1) showed no binding to the native proteins. For biodistribution studies, L-VHH and D-VHH with single radioactive indium diethylenetriamine-pentaacid ($^{111}$In-DTPA) labeling at the C-terminus were designed and synthesized by the established protocol. The distribution profiles were essentially similar between L-VHH and D-VHH, in which the probes accumulated in the kidney within 15 min after intravenous administration in mice because of the small molecular size of VHHs. Comparative assessment of the immunogenicity responses revealed that D-VHH-induced levels of anti-drug antibody (ADA) generation were significantly lower than those of native VHH, regardless of the peptide sequences and administration routes. The resulting scaffold investigated should be applicable in the design of D-VHHs with various C-terminal CDR3 sequences, which can be identified by screening using display technologies.*

The variable domain of the heavy chain of the heavy chain antibodies (VHHs), also referred to as nanobodies, are single-domain antibody fragments derived from heavy chain antibodies found in *Camelidae*.[1] VHHs consist of four frame regions (FR1–4) and three complementarity-determining regions (CDR1–3) (Figure 1).[2] Among the three CDRs, the most variable CDR3 can contribute to the potent and selective interaction with an antigen by arranging the length and amino acid usage.[3] VHHs are attractive alternatives to conventional antibodies in biotechnology and pharmaceuticals due to their favorable properties, such as small size, high stability, and improved solubility.[4] In fact, the anti-von Willebrand factor (vWF) bivalent VHH, caplacizumab, was approved as the

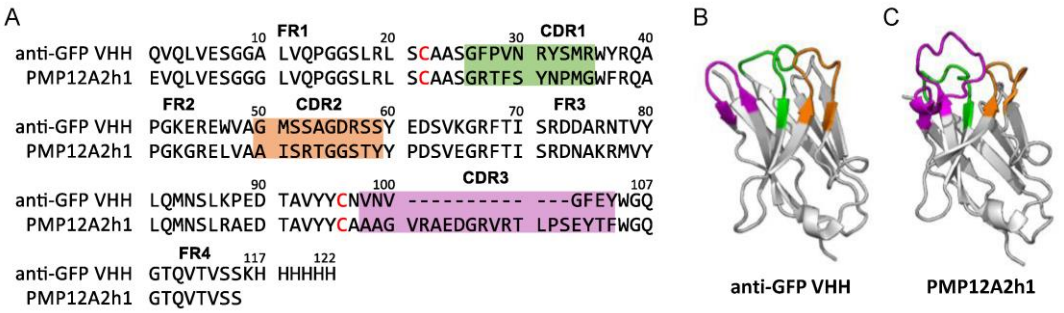


**Figure 1.** Sequences and structures of representative VHHs. (A) Sequences of anti-GFP VHH and PMP12A2h1. Red-colored letters: cysteines; CDR (highlighted areas): complementarity-determining regions; FR: framework regions. (B, C) Structures of anti-GFP VHH (B) or PMP12A2h1 (C). CDRs are colored with green (CDR1), orange (CDR2), or magenta (CDR3).

first VHH-based drug in 2018 for the treatment of acquired thrombotic thrombocytopenic purpura.[5] Subsequently, the anti-tumor necrosis factor (TNF)/anti-human serum albumin (HSA) bispecific VHH ozoralizumab received its approval in 2022 for the treatment of rheumatoid arthritis.[6] Although VHHs have relatively low immunogenicity, VHH-based therapeutics still induce anti-drug antibody (ADA) production.[7] Indeed, a phase III trial of ozoralizumab showed that ADA production was increased in approximately 40% of ozoralizumab-treated patients.[8]

As described in Preface, the author hypothesized that a mirror-image form of VHH (D-VHH) can be a novel D-protein-based scaffold with high-affinity target binding and low immunogenicity for diagnostic and therapeutic purposes. For mirror-image screening to explore D-VHHs with the expected target binding, several VHH libraries for phage display,[9] yeast display[10] and mRNA/cDNA display[11] technologies are available. Another essential technology for developing D-VHH-based therapeutic agents is the chemical synthesis of a D-VHH scaffold with a hit sequence obtained from mirror-image screening. In this study, the author aimed to establish a practical synthetic route to D-VHH scaffolds, that is applicable to various possible bioactive sequences from commercially available reagents. Comparative investigations of the biodistribution and immunogenicity of the resulting D-VHH in mice were performed to characterize the biological properties of D-VHHs.

**Chemical Synthesis of Anti-GFP VHH**. The author designed a synthetic scheme for anti-GFP VHH[12,13] as a case study using a sequential native chemical ligation (NCL) strategy.[14] Anti-GFP VHH was identified by immunization of *Camelidae* with green fluorescent protein (GFP) followed by phage display screening.[12] To establish a common

protocol for preparing VHHs with variable CDR3 sequences, ligation of the C-terminal CDR3 segment in the final step is favorable. The author envisioned that two consensus cysteines in VHH frame regions ($Cys^{22}$ in FR1 and $Cys^{96}$ in FR3) should be suitable for NCLs. Based on these structural characteristics and application perspectives, the author designed three peptide segments L-**1**–L-**3** with the appropriate accessory group(s) for N-to-C sequential ligations (Figure 2): the N-terminal thioester segment L-**1** [anti-GFP $VHH^{1–21}$]; the middle segment L-**2** [anti-GFP $VHH^{22–95}$] with an activable first-generation Dawson linker (Dbz) at the C-terminus; and the C-terminal segment L-**3** [anti-GFP $VHH^{96–122}$] with a histidine tag.[13]

Initially, the author investigated the synthesis of three segment peptides by standard Fmoc-based solid-phase peptide synthesis (SPPS) using 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU)/1-hydroxybenzotriazole (HOBt)/$(i\text{-Pr})_2$NEt activation (Figure 3).[15] For the preparation of N-terminal peptide thioester L-**1**, the protected peptide resin was synthesized with a Dbz moiety,[16] which was converted to the benzimidazol-2-one (Nbz) form by treatment with 4-nitrophenyl chloroformate and $(i\text{-Pr})_2$NEt. Global deprotection and cleavage from the resin followed by reaction with sodium 2-mercaptethanesulfonate (MESNa) provided thioester L-**1**. Similarly, the C-terminal segment L-**3** was obtained using Rink amide resin in satisfactory yields. In contrast, the yield and purity of middle segment L-**2** (74 residues) by solid-phase synthesis were low. Thus, an alternative strategy was designed to synthesize L-**2**. The author focused on $Ala^{54}$ in the sequence of L-**2**, which can be obtained by desulfurization of Cys.[17,18] With the intent to employ this $Ala^{54}$ as an additional NCL site, peptide L-**2** was split into two peptide segments: [$Cys(Acm)^{22}$]-anti-GFP $VHH^{22–53}$ (L-**4**) and [$Cys^{54}$]-anti-GFP $VHH^{54–95}$ (L-**5**). Peptide thioester L-**4** with an acetamidomethyl (Acm)-protected Cys at the N-terminus was synthesized by using the same procedure for the synthesis of L-**1**. Peptide L-**5** with Dbz and Arg accessory groups at the C-terminus was synthesized by SPPS on a Rink amide-PEG resin. Peptide thioester L-**4** and cysteine peptide L-**5** were subjected to 4-mercaptophenylacetic acid (MPAA)-mediated NCL conditions to provide peptide L-**6**.[19] The radical-mediated metal-free desulfurization (MFD) of L-**6** with VA-044[17] followed by deprotection of the Acm group on the N-terminal Cys side-chain in L-**7** with $PdCl_2$[20] afforded the middle peptide segment L-**2**.

L-**1** $^{1}$QVQLVESGGALVQPGGSLRLS-MES

L-**2** $^{22}$CAASGFPVNRYSMRWYRQAPGKEREWVAGMSS$^{54}$AGDRSSYEDSVKGRFTISRDDARNTVYLQMNSLKPEDTAVYY-Dbz-R-$NH_2$

L-**3** $^{96}$CNVNVGFEYWGQGTQVTVSSKHHHHHH-$NH_2$

**Figure 2.** Amino acid sequences of the initially planned peptide segments for anti-GFP VHH.

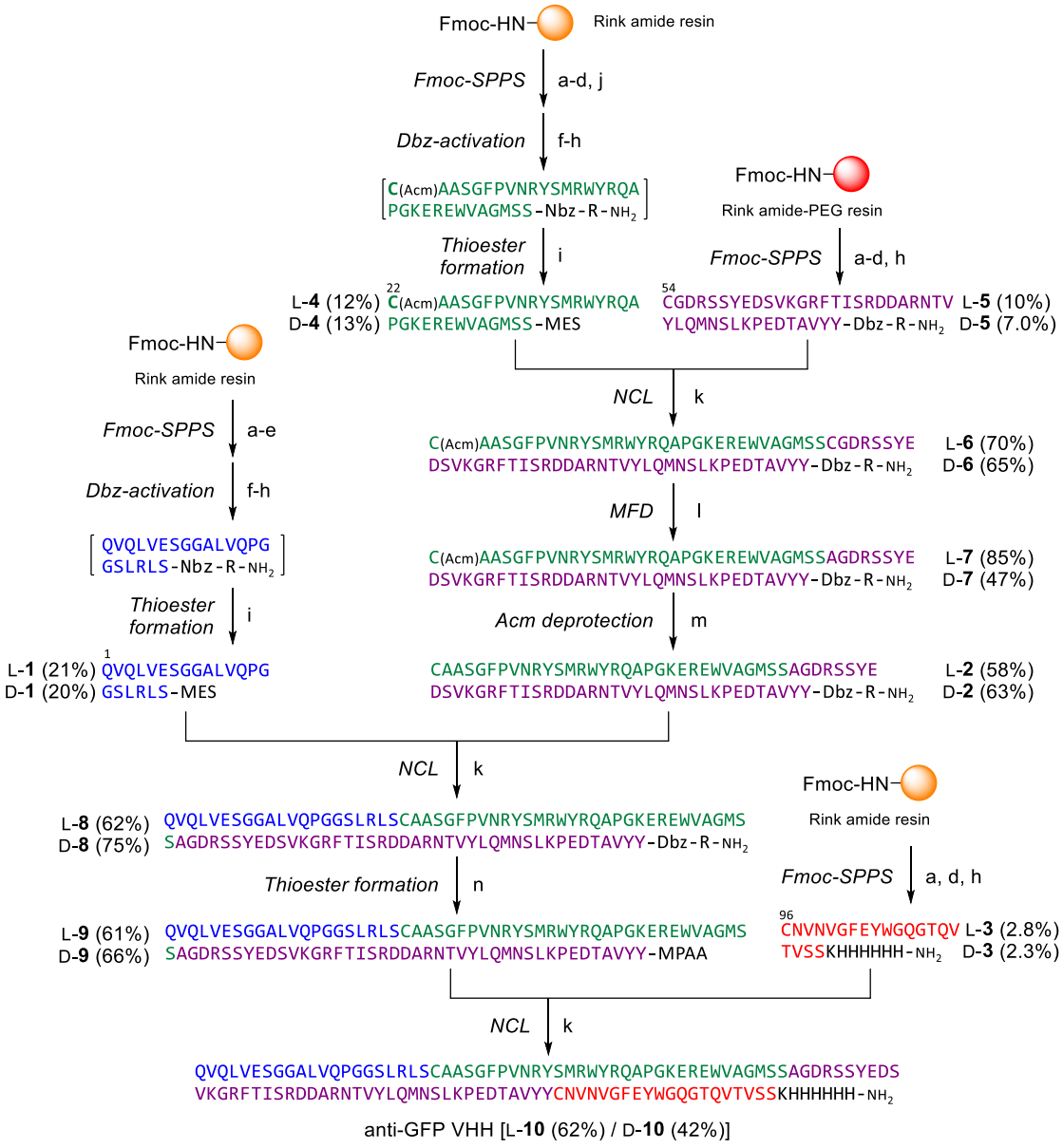


**Figure 3.** Synthesis of L- and D-anti-GFP VHH. *Reagents and conditions*: (a) 20% piperidine/DMF; (b) Fmoc-Arg(Pbf)-OH, HBTU, HOBt, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF; (c) Fmoc-Dbz-OH, HBTU, HOBt, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF; (d) Fmoc-Xaa-OH, HBTU, HOBt, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF [repeat (d) until the sequence was completed]; (e) $Boc_2O$, (*i*-Pr)$_2$NEt, and DMF; (f) 4-nitrophenyl chloroformate and $CH_2Cl_2$; (g) (*i*-Pr)$_2$NEt and DMF; (h) TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (EDT) (80:5:5:5:5); (i) MESNa, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 6.0); (j) Boc-Cys(Acm)-OH, HBTU, HOBt, (*i*-Pr)$_2$NEt, and DMF; (k) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (l) VA-044, MESNa, TCEP, 6 M guanidine·HCl, and 100 mM phosphate buffer (pH 6.5); (m) $PdCl_2$, 6 M guanidine·HCl, 100 mM phosphate buffer, and then DTT; and (n) $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, then MPAA and TCEP.

With the originally planned peptide segments (L-**1**–L-**3**) to hand, the author investigated the N-to-C sequential ligations. Peptide thioester L-**1** and middle segment L-**2** were ligated in the presence of MPAA to give L-anti-GFP VHH$^{1–95}$ (L-**8**) in 62% yield.[19] $NaNO_2$-mediated activation[21] of the Dbz linker in L-**8** followed by treatment with MPAA provided thioester L-**9**. Of note, the conversion of L-**8** into the MES thioester failed because of the difficulty in separating the desired product from byproducts and reagents. The subsequent NCL between thioester L-**9** and the C-terminal segment L-**3** afforded the desired peptide L-**10** (L-anti-GFP VHH). Notably, during the author's investigation, two synthetic studies of VHH were reported,[22,23] in which a solubility tag or pseudoproline dipeptides were used to improve the solubility of the peptide segments and/or the purification process of the NCL products. In this synthetic process, taking advantage of the finely tuned NCL process and peptide segments designed, the author achieved chemical synthesis of anti-GFP VHH from commercially available reagents without using auxiliary functional groups. The mirror-image protein (D-anti-GFP VHH, D-**10**) was also synthesized using D-amino acids, glycine, and the same procedure.

The resulting synthetic anti-GFP VHH was subjected to refolding conditions to obtain the functional protein. After lyophilized L-anti-GFP VHH (L-**10**) was denatured in 6 M guanidine containing 40 mM DTT, the solution was diluted 100-fold in PBS (pH 7.4). The solution was concentrated by ultrafiltration to ~1.0 mg/mL. Synthetic D-anti-GFP VHH (D-**10**) was folded in the same manner. Circular dichroism (CD) spectra of synthetic L- and D-anti-GFP VHH were recorded to confirm the structures (Figure 4). The CD spectrum of synthetic L-anti-GFP VHH was consistent with recombinant L-anti-GFP VHH,[24] with a negative band at ~210 to 220 nm derived from the predominant β-sheet structure. Symmetrical spectra of L-anti-GFP VHH and D-anti-GFP VHH were observed, suggesting that both synthetic L- and D-anti-GFP VHH were folded correctly.

The function of folded synthetic L-anti-GFP VHH was examined by evaluating the binding to enhanced green fluorescent protein (EGFP) using several methods. As a

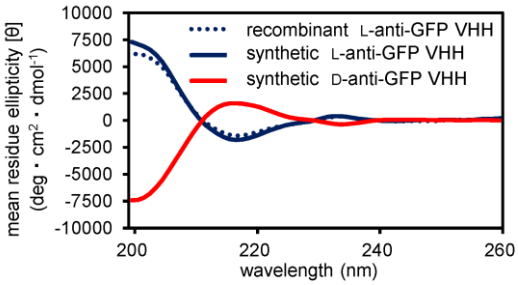


**Figure 4.** CD spectra of recombinant L-anti-GFP VHH, synthetic L-anti-GFP VHH and D-anti-GFP VHH.

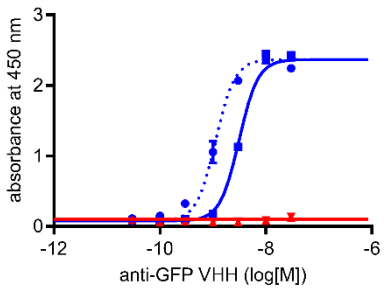


**Figure 5.** Binding curves of L-anti-GFP VHH with EGFP (blue) and mCherry (red). The binding was determined using enzyme-linked immunosorbent assay (ELISA). Synthetic L-anti-GFP VHH (solid line) and recombinant L-anti-GFP VHH (dotted line). Mean absorbance values ± SD are shown.

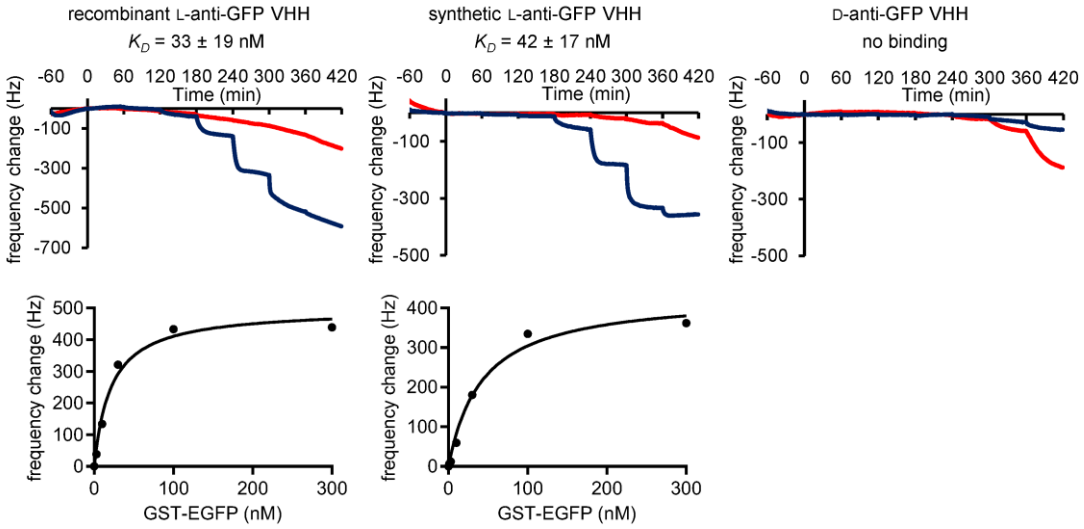


**Figure 6.** Representative data of the quartz crystal microbalance (QCM) analysis of recombinant L-anti-GFP VHH, synthetic L-anti-GFP VHH, and D-anti-GFP VHH with EGFP (blue) and mCherry (red). A Langmuir model was used for determining binding affinities from triplicate assays.

negative control, mCherry was employed, which is not recognized by anti-GFP VHH.[13] In an enzyme-linked immunosorbent assay (ELISA), the binding of synthetic L-anti-GFP VHH with EGFP was observed, and binding was similar to that of recombinant L-anti-GFP VHH with EGFP (Figure 5). In quartz crystal microbalance (QCM) analysis,[25] the binding affinity of synthetic L-anti-GFP VHH with EGFP ($K_D$ = 42 ± 17 nM) was similar to that obtained for recombinant L-anti-GFP VHH ($K_D$ = 33 ± 19 nM) (Figure 6). In contrast, the binding of D-anti-GFP VHH with EGFP was not detected.[26]

**Chemical Synthesis of PMP12A2h1**. The author next investigated the applicability of the established synthetic protocol of VHHs. PMP12A2h1 is a monomeric VHH of caplacizumab, which binds to the A1 domain of vWF.[5] PMP12A2h1 was identified by immunization of a llama with recombinant vWF followed by humanization.[27] Because caplacizumab is an approved therapeutic agent, the monomeric PMP12A2h1 framework represents a promising mirror-image VHH scaffold with favorable pharmaceutical properties, including protein stability and physicochemical properties for drug discovery. The author planned the chemical synthesis of PMP12A2h1 via NCLs at two conserved cysteines ($Cys^{22}$ and $Cys^{96}$) using a similar protocol for synthesizing anti-GFP VHH. For this purpose, three peptide segments L-**11**–L-**13** were designed (Figure 7): the N-terminal segment L-**11** [PMP12A2h1$^{1–21}$]; the middle segment L-**12** [PMP12A2h1$^{22–95}$] with an activable Dbz moiety at the C-terminus; and the C-terminal segment L-**13** [PMP12A2h1$^{96–128}$]. To synthesize the middle segment L-**12** (74 residues), the author used $Ala^{50}$ as an additional NCL site, which can be constructed by desulfurization of $Cys^{50}$ after NCL between PMP12A2h1$^{22–49}$ (L-**14**) and [$Cys^{50}$]-PMP12A2h1$^{50–95}$ (L-**15**).[18]

Peptides L-**11** and L-**14** with activable Nbz and MeNbz[28] moieties, respectively, at the C-terminus, were synthesized by microwave-assisted Fmoc-SPPS using DIC/Oxyma activation.[29] The N-terminal segment L-**11** was obtained in a yield comparable to that of peptide L-**1** for anti-GFP VHH, in which two amino acids in the framework are substituted. In contrast, the yield of the second segment L-**14** containing CDR1 was better than that of the corresponding L-**4**, presumably because of the shorter length and/or amino acid compositions of CDR1. Peptides L-**13** and L-**15** containing Asp-X sequences (X = Gly, Ser, Asn) were synthesized by standard Fmoc-SPPS to prevent aspartimide formation.[30] The solubility and purification efficiency of the C-terminal segments, such as L-**3** (for anti-GFP VHH) and L-**13** (for PMP12A2h1), were highly dependent on the CDR3 sequence. Peptide L-**13** with three arginine residues in the long CDR3 was obtained in moderate yield, whereas the yield of the corresponding L-**3** with no basic residues in the CDR3 was very low. Peptide L-**15** with the CDR2 sequence was obtained in a yield lower than that of the corresponding L-**5**, although the solubility of these peptides was similar.

MPAA-mediated NCL between MeNbz-peptide L-**14** and cysteine peptide L-**15** provided peptide L-**16**. VA-044-mediated desulfurization of L-**16** followed by $PdCl_2$-mediated Acm deprotection on the N-terminal Cys side-chain in L-**17** afforded the middle peptide segment L-**12**. Nbz-peptide L-**11** and middle segment L-**12** were ligated in the presence of MPAA to give L-PMP12A2h1$^{1–95}$ (L-**18**) in 64% yield.[19] Note that to optimize the preparation of peptide L-**17**, the author investigated further splitting and NCL at $Ala^{75}$ and newly designed two peptide segments (L-**22** and L-**23**); however, the overall yield

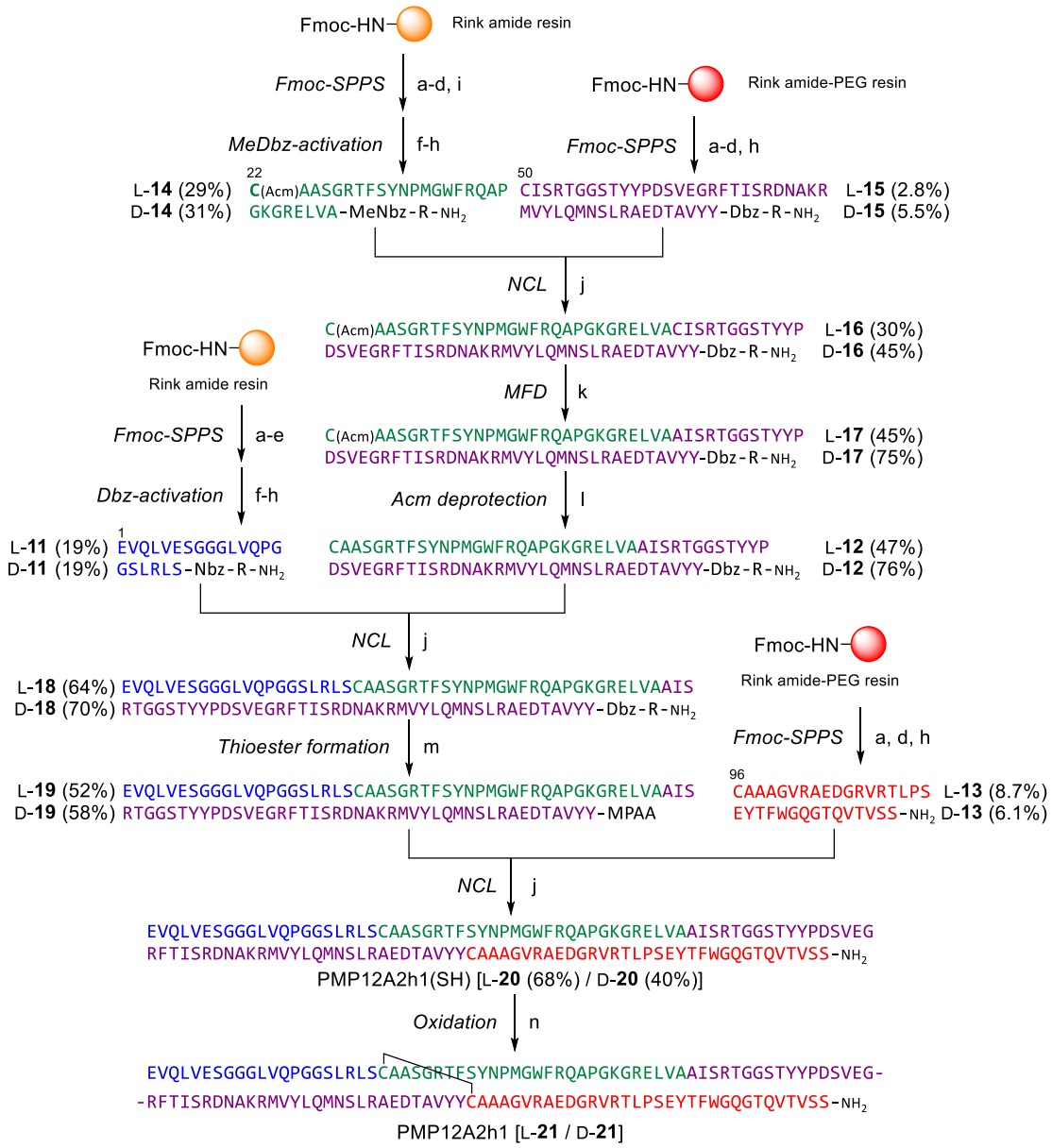


**Figure 7.** Synthesis of L- and D-PMP12A2h1. *Reagents and conditions*: (a) 20% piperidine/DMF; (b) Fmoc-Arg(Pbf)-OH, Oxyma Pure, DIC, DMF, and then 20% piperidine/DMF; (c) Fmoc-Dbz-OH or Fmoc-MeDbz-OH, HBTU, HOBt, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF; (d) Fmoc-Xaa-OH, Oxyma Pure, DIC, DMF, and then 20% piperidine/DMF [repeat (d) until the sequence was completed]; (e) $Boc_2O$, (*i*-Pr)$_2$NEt, and DMF; (f) 4-nitrophenyl chloroformate and $CH_2Cl_2$; (g) (*i*-Pr)$_2$NEt and DMF; (h) TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5); (i) Boc-Cys(Acm)-OH, Oxyma Pure, DIC, and DMF; (j) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (k) VA-044, TCEP, MESNa, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 6.5); (l) $PdCl_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, and then DTT; (m) $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, then MPAA and TCEP; (n) 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB).

was unexpectedly not improved (Figure 8). Activation of the Dbz linker in L-**18** using $NaNO_2$ followed by thioester formation provided peptide L-**19**.

The final NCL between thioester L-**19** and C-terminal segment L-**13** gave the expected full-length L-PMP12A2h1 (SH) (L-**20**). The synthetic protein was subjected to refolding conditions in PBS to yield the bioactive protein. Subsequently, oxidation by treatment with 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB)[31] provided the disulfide form of PMP12A2h1 (L-**21**, native form). The mirror-image protein [D-PMP12A2h1 (D-**21**)] was prepared by the same procedure.

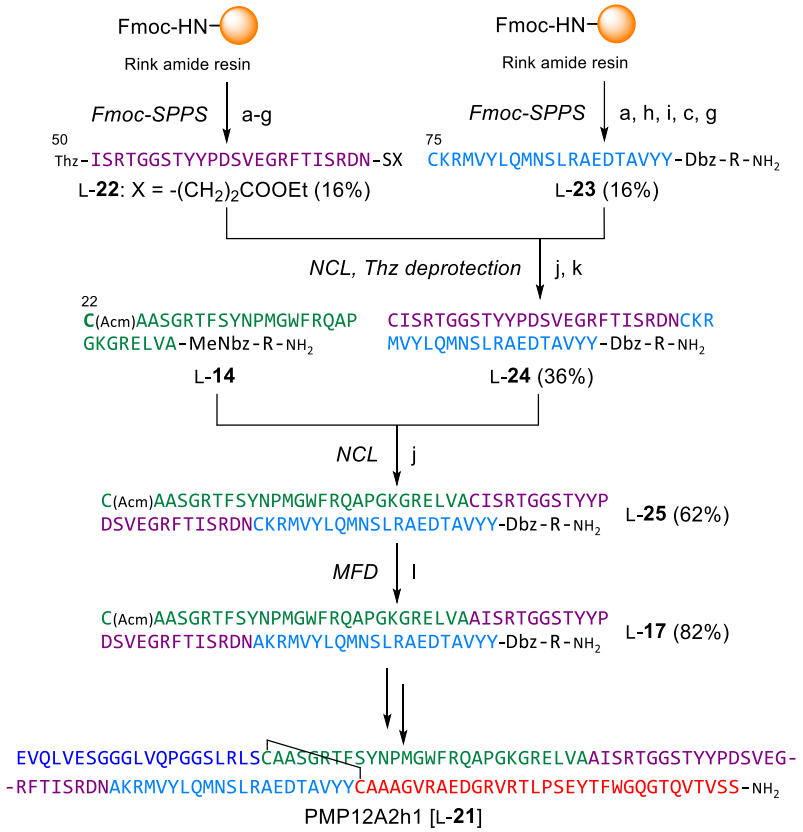


**Figure 8.** Synthesis of L-PMP12A2h1 using five peptide segments. *Reagents and conditions*: (a) 20% piperidine/DMF; (b) Fmoc-Asp-OAllyl, Oxyma Pure, DIC, DMF, and then 20% piperidine/DMF; (c) Fmoc-Xaa-OH, HBTU, HOBt, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF [repeat (c) until the sequence was completed]; (d) $Boc_2O$, (*i*-Pr)$_2$NEt, and DMF; (e) $Pd(PPh_3)_4$, $PhSiH_3$, and $CH_2Cl_2$; (f) ethyl 3-mercaptopropionate, HOAt, DIC, (*i*-Pr)$_2$NEt, $CH_2Cl_2$, and DMF; (g) TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5); (h) Fmoc-Arg(Pbf)-OH, HBTU, HOBt, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF; (i) Fmoc-Dbz-OH, HBTU, HOBt, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF; (j) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (k) methoxyamine; (l) VA-044, TCEP, MESNa, 6 M guanidine·HCl, and 100 mM phosphate buffer (pH 6.5).

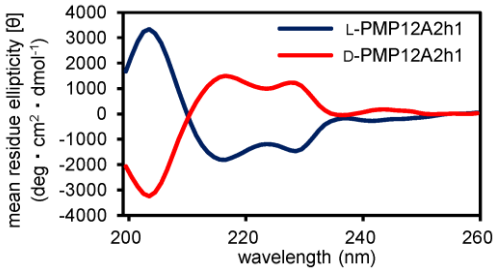


**Figure 9.** CD spectra of L- and D-PMP12A2h1.

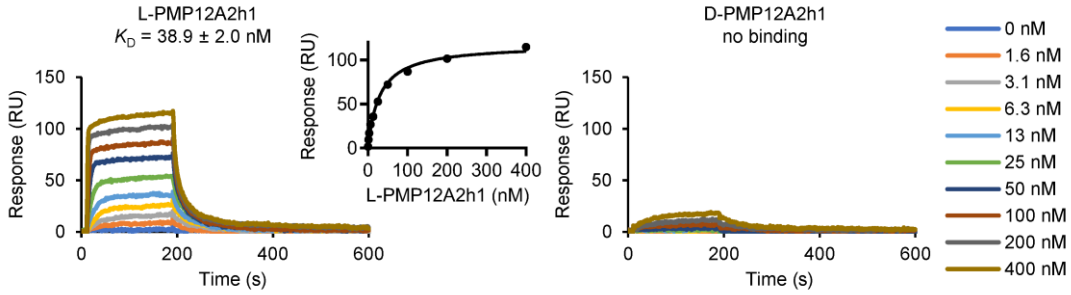


**Figure 10.** Representative SPR data analysis for the binding of folded L- and D-PMP12A2h1 to the vWF-A1 domain. A Langmuir model was used for determining binding affinities from triplicate assays.

The CD spectrum of L-PMP12A2h1 was different from that of anti-GFP VHH, which was attributed to the different sequences in the framework region and CDR.[32] The CD spectra of L-PMP12A2h1 and D-PMP12A2h1 were symmetrical (Figure 9), suggesting that folded L-PMP12A2h1 and D-PMP12A2h1 had enantiomeric conformations. Surface plasmon resonance (SPR) analysis revealed that L-PMP12A2h1 bound to the vWF-A1 domain with sufficient binding affinity [$K_D$ = 38.9 ± 2.0 nM], whereas no binding signal was observed between D-PMP12A2h1 and the vWF-A1 domain (Figure 10).[26] These results indicate that the synthetic PMP12A2h1 reproduced the desired function.

The author comparatively investigated the properties of the reduced forms [L-PMP12A2h1 (SH) and D-PMP12A2h1 (SH)] compared with those of the oxidized forms [L-PMP12A2h1 and D-PMP12A2h1]. In the CD spectral analysis, no significant difference between the reduced and oxidized forms was observed (Figures 9 and 11). The affinity of L-PMP12A2h1 (SH) to the vWF-A1 domain by SPR analysis [$K_D$ = 39.5 ± 3.2 nM] was identical to that of L-PMP12A2h1 (Table 1). In contrast, the thermal stability differed between that of the reduced and oxidized forms. In the denaturing analysis

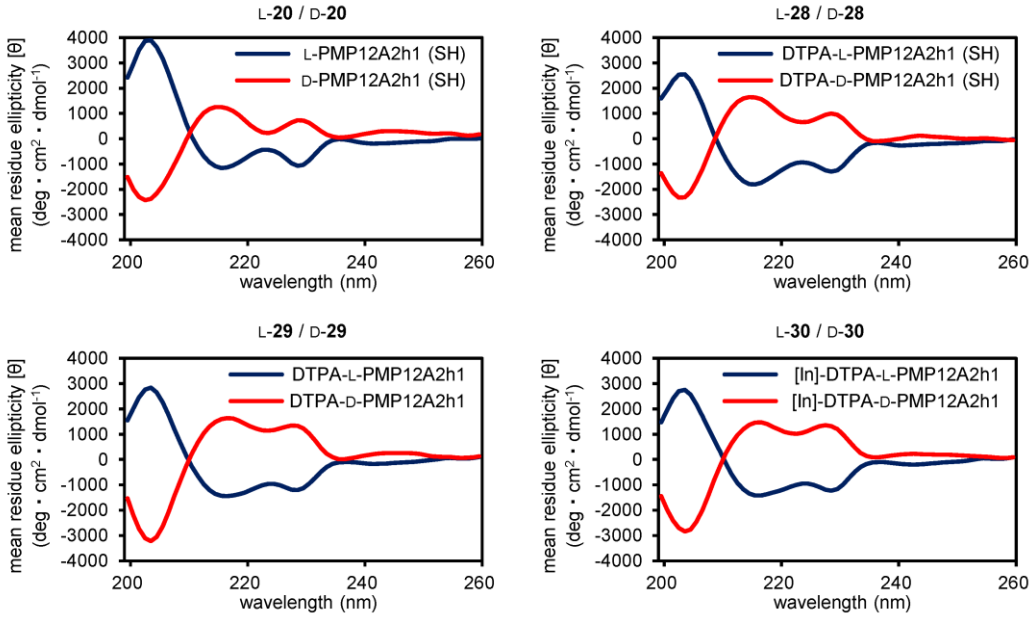


**Figure 11.** CD spectra of synthetic L- and D- PMP12A2h1 derivatives.

**Table 1.** SPR analysis of synthetic PMP12A2h1 derivatives binding with vWF-A1 domain.

| Analyte | $K_D$ (nM)[a] |
|---|---|
| L-PMP12A2h1 (SH) (L-**20**) | 39.5 ± 3.2 |
| DTPA-L-PMP12A2h1 (SH) (L-**28**) | 46.6 ± 0.1 |
| DTPA-L-PMP12A2h1 (L-**29**) | 52.9 ± 2.0 |
| [In]-DTPA-L-PMP12A2h1 (L-**30**) | 53.8 ± 2.2 |
| D-PMP12A2h1 (SH) (D-**20**) | No binding |
| DTPA-D-PMP12A2h1 (SH) (D-**28**) | No binding |
| DTPA-D-PMP12A2h1 (D-**29**) | No binding |
| [In]-DTPA-D-PMP12A2h1 (D-**30**) | No binding |

[a] $K_D$ values were determined from triplicate assays.

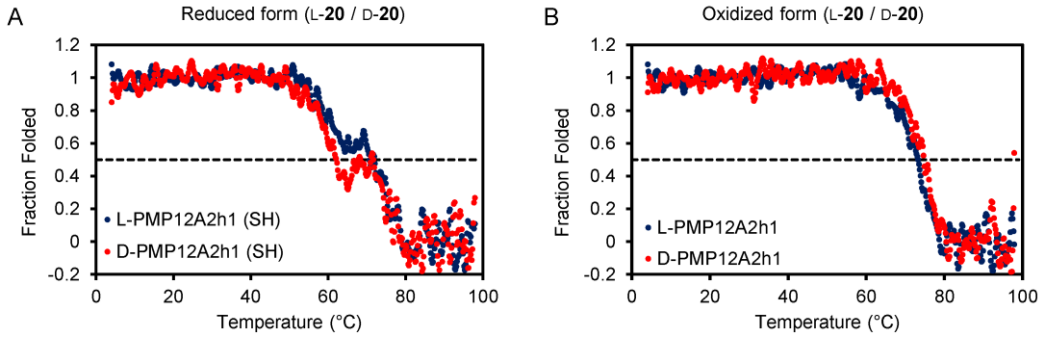


**Figure 12.** Thermal stability of synthetic L- and D- PMP12A2h1 derivatives measured by CD at 203 nm. (A) L-PMP12A2h1 (SH form) (blue) and D-PMP12A2h1 (SH form) (red). (B) L-PMP12A2h1 (blue) and D-PMP12A2h1 (red).

monitored by measuring the CD signal at 203 nm, the temperature-dependent transition of PMP12A2h1 (SH) began at ~50 °C, whereas denaturation of the oxidized form started above 60 °C (Figure 12). These observations correspond to the general properties of the disulfide bond between two cysteines in VHH, which is not indispensable for high-affinity binding to antigens but contributes to thermal stability.[33]

**Comparative Biodistribution Study of L-VHH and D-VHH.** There have been several reports on the pharmacokinetic profiles of L-VHHs, including PMP12A2h1. The small molecular size (approximately 15 kDa) of VHHs has been shown to shorten the half-life circulation via glomerular filtration in kidneys.[27,34–36] This rapid clearance from circulation and nontarget organs and good tissue penetration are favorable biodistribution properties for molecular imaging.[36,37] However, there has been no report on the biodistributions of mirror-image protein scaffolds such as D-VHH. For pharmaceutical applications of the D-VHH scaffold, unexpected accumulation(s) of D-VHH in a specific organ(s) would be unfavorable. Thus, investigating the biodistribution of synthetic D-VHH scaffolds is worthwhile. PMP12A2h1 is recognized by human vWF but not by mouse vWF, suggesting no specific target molecules of L-PMP12A2h1 and D-PMP12A2h1 in mice.[27] Therefore, the author assumed that a study using L-PMP12A2h1 and D-PMP12A2h1 in mice is an appropriate model to assess the biodistribution of a pair of mirror-image VHHs.

For this biodistribution study, the author designed and synthesized a PMP12A2h1 probe labeled with radioactive indium diethylenetriamine-pentaacid ($^{111}$In-DTPA) (Figure 13A). Based on the structural information on VHHs, the author planned the DTPA conjugation at the C-terminus via an aminocaproic acid (Ahx) linker to prevent the loss of PMP12A2h1 activity. The DTPA-labeled C-segment (L-**27**) was prepared by constructing the protected DTPA moiety on the Lys side-chain on a solid support using the DTPA precursor **26**.[38] NCL between peptide L-**19** and C-segment L-**27** provided DTPA-labeled L-PMP12A2h1 (SH) (L-**28**). The disulfide form L-**29** was synthesized by using the same procedure for synthesizing L-**21**. Peptide L-**29** was readily labeled with $^{111}$In under acidic conditions to give $^{111}$In-labeled L-PMP12A2h1 (L-**30**) with sufficient radiochemical purity after purification by ultrafiltration to remove unreacted $^{111}InCl_3$ (Figure 13B). $^{111}$In-DTPA-labeled D-PMP12A2h1 (D-**30**) was similarly synthesized, in which the DTPA precursor **26** was attached to the D-Lys side-chain on a solid support.

The structure and function of the resulting labeled L-PMP12A2h1 derivatives L-**28**–L-**30** were evaluated. CD spectral analysis suggested that the secondary structures were conserved between the intact protein L-**21** and DTPA-labeled proteins L-**28**–L-**30** (Figure

11). The labeled D-PMP12A2h1 derivatives D-**28**–D-**30** had enantiomeric structures of the L-congener. SPR measurements confirmed that the biological activity of labeled PMP12A2h1 derivatives to the target vWF-A1 domain was reproducible [$K_D$ (L-**28**) = 46.6 ± 0.1 nM; $K_D$ (L-**29**) = 52.9 ± 2.0 nM; $K_D$ (L-**30**) = 53.8 ± 2.2 nM] (Table 1). No binding of the mirror-image PMP12A2h1 derivatives (D-**28**–D-**30**) to native vWF-A1 domain was observed.[26] These results suggest that DTPA labeling and In(III) chelating did not affect the structures and bioactivity of unlabeled PMP12A2h1 proteins.

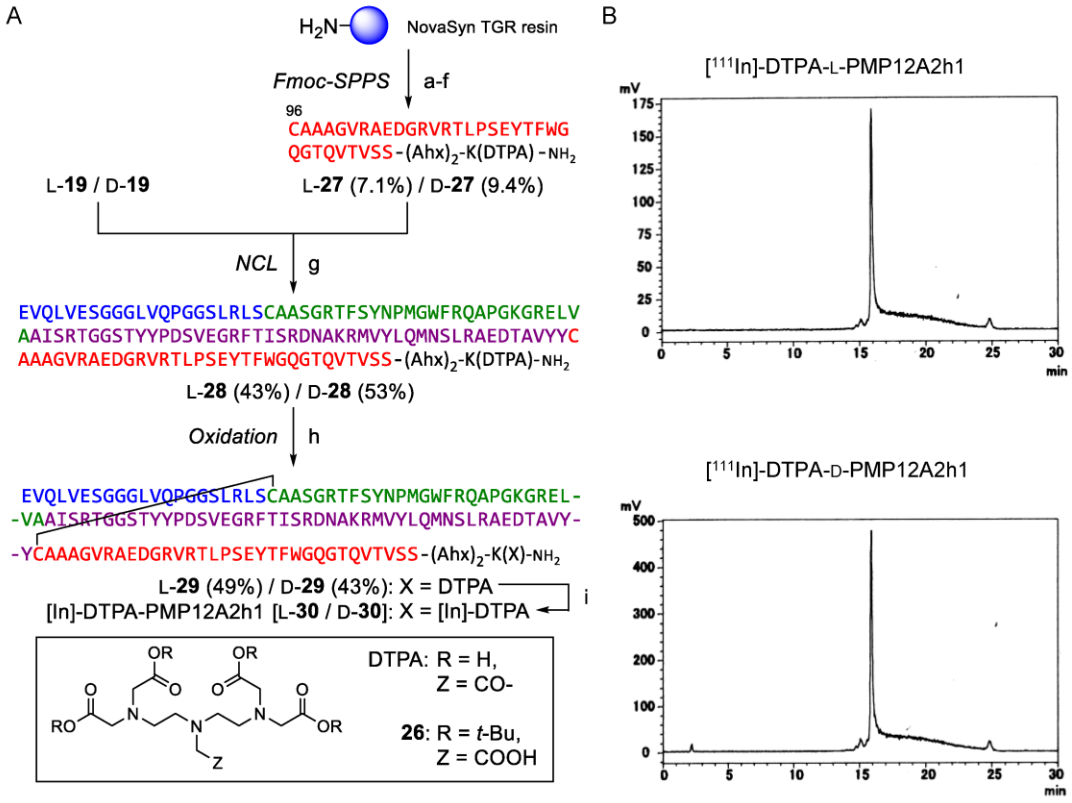


**Figure 13.** Preparation of [$^{111}$In]-labeled PMP12A2h1. (A) Synthesis of [$^{111}$In]-DTPA-L-PMP12A2h1 and [$^{111}$In]-DTPA-D-PMP12A2h1. *Reagents and conditions*: (a) Fmoc-Lys(Mtt)-OH, HATU, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF; (b) $CH_2Cl_2$/TFA/TIS = 94:1:5; (c) **26**, HATU, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF; (d) Fmoc-Ahx-OH, HATU, (*i*-Pr)$_2$NEt, DMF, and then 20% piperidine/DMF [repeat (d)]; (e) Fmoc-Xaa-OH, Oxyma Pure, DIC, and then 20% piperidine/DMF [repeat (e) until the sequence was completed]; (f) TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5); (g) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (h) DTNB; (i) $InCl_3$. (B) Radiochemical Purity of [$^{111}$In]-DTPA-L-PMP12A2h1 and [$^{111}$In]-DTPA-D-PMP12A2h1. *HPLC conditions:* Cosmosil Protein-R column (4.6 × 150 mm, Nacalai Tesque), linear gradient of 25-45% $CH_3CN$ containing 0.05% TFA at a flow rate of 1 mL/min over 20 min.

For biodistribution studies, male ddY mice received intravenous administrations of $^{111}$In-labeled L-PMP12A2h1 (L-**30**) or $^{111}$In-DTPA-labeled D-PMP12A2h1 (D-**30**). Among the 13 organs evaluated, a slightly higher accumulation of D-VHH in the stomach than that of L-VHH was observed until 2 h after administration for some mice (Figure 14A). Although the reason for this gastric accumulation is unclear, it may arise from the unique properties of D-VHH, which is of interest for the potential use of D-VHH-based therapeutic agents for gastric disorders. In the other organs, D-VHH showed a similar distribution to L-VHH (Figure 14B–D).[39] Rapid renal excretion of L-VHH and D-VHH was observed, suggesting that the predominant renal clearance of VHH-based agents persisted even for D-VHH.

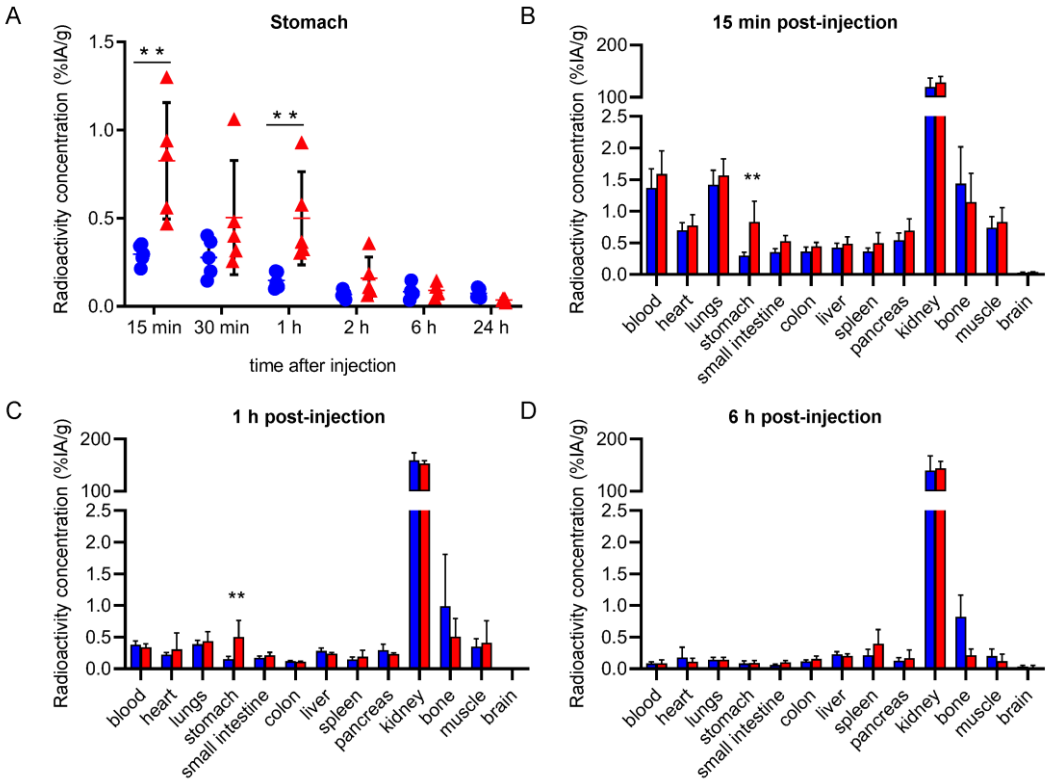


**Figure 14.** Biodistribution of [$^{111}$In]-DTPA-L-PMP12A2h1 (blue) and [$^{111}$In]-DTPA-D-PMP12A2h1 (red) in mice. Error bars indicate standard deviations of the mean percentage of injected dose per gram. Differences between [$^{111}$In]-DTPA-L-PMP12A2h1 and [$^{111}$In]-DTPA-D-PMP12A2h1 were analyzed by the unpaired Mann–Whitney U test. $^{**}p < 0.01$. (A) The radioactivity of the stomach at given time points (15 min, 30 min, 1 h, 2 h, 6 h, 24 h) after intravenous administration ($N$ = 5). (B–D) The radioactivity of the 13 organs measured at 15 min (B), 1 h (C), and 6 h (D) after intravenous administration ($N$ = 5).

**Immunogenicity of L-VHH and D-VHH.** Quantitative analysis of ADA production by administering both anti-GFP VHH and PMP12A2h1 enantiomers was performed to evaluate the immunogenicity of the synthetic L-VHHs and D-VHHs. Serum was collected GFP VHH-treated group, L-anti-GFP VHH-binding antibodies emerged on days 21 and

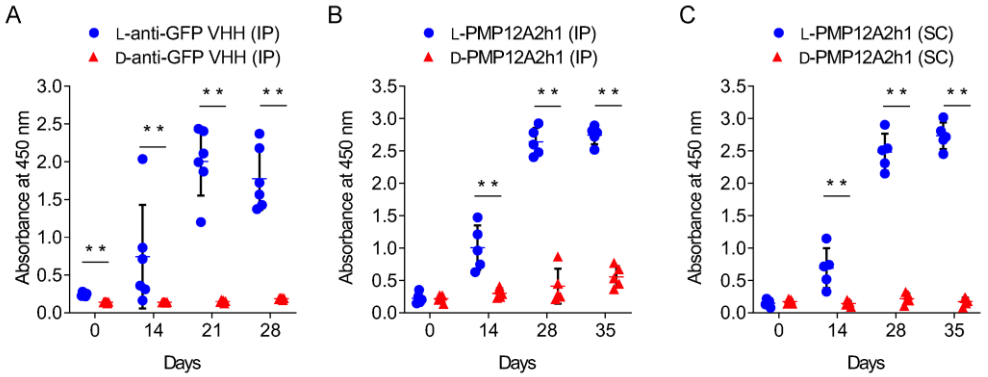


**Figure 15.** Immunogenicity of L- and D-VHHs. (A) ADA generation in mouse sera on days 0, 14, 21, and 28 after intraperitoneal injection of L- or D-anti-GFP VHH (50 μg/injection) on days 0, 14, and 21. The concentration of the coated antigen (L- or D-anti-GFP VHH) for ELISA is 1000 ng/mL. (B, C) ADA generation in mouse sera on days 0, 14, 28, and 35 after intraperitoneal (B) or subcutaneous (C) injection of L- or D-PMP12A2h1 (25 μg/injection) on days 0, 14, and 28. The concentration of coated antigen (L- or D-PMP12A2h1) for ELISA is 1000 ng/mL. Statistics were carried out using the Mann–Whitney U test between two groups. $^{**}p < 0.01$.

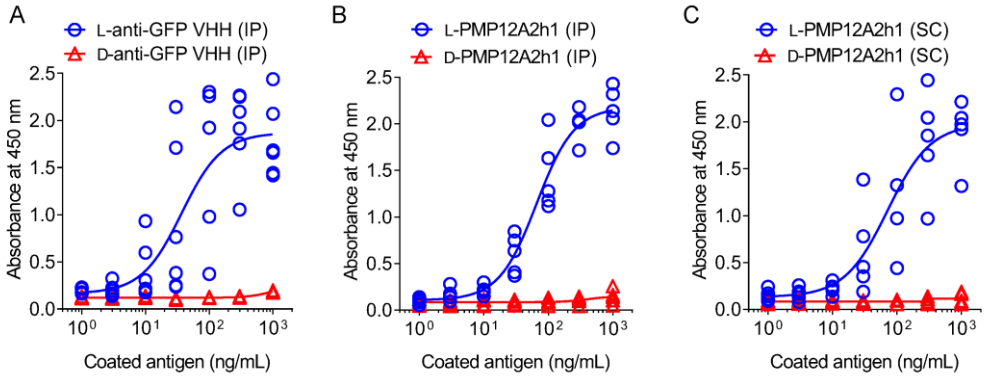


**Figure 16.** Quantitative analysis of ADA generation against synthetic L-VHH and D-VHH. (A) Detection of ADA in mouse sera (day 28) from intraperitoneal injection of L-anti-GFP VHH or D-anti-GFP VHH (50 μg/injection) by ELISA with varying concentrations of the plate-coating antigen (anti-GFP VHH). (B) Detection of ADA in mouse sera (day 35) after intraperitoneal injection of L-PMP12A2h1 or D-PMP12A2h1 (25 μg/injection) by ELISA with varying concentrations of the plate-coating antigen (PMP12A2h1). (C) Detection of ADA in mouse sera (day 35) after subcutaneous injection of L-PMP12A2h1 or D-PMP12A2h1 (25 μg/injection) by ELISA with varying concentrations of the plate-coating (PMP12A2h1).

28 after initial immunization by intraperitoneal administration (Figures 15A and 16A). Even on day 14, a significant increase in the ADA level was observed in some mice. In contrast, in the D-anti-GFP VHH-treated group, no generation of D-anti-GFP VHH-binding antibodies was observed over the entire observation period of the experiment.

ADA production by L-PMP12A2h1 or D-PMP12A2h1 was also comparatively investigated. Intraperitoneal administration of L-PMP12A2h1 induced a time-dependent increase in ADA levels for L-PMP12A2h1, whereas administration of D-PMP12A2h1 yielded a significantly lower level of ADA production (Figures 15B and 16B). The author also assessed the effect of the route of administration of L-VHH and D-VHH on ADA production (Figures 15C and 16C). Subcutaneous injection of L-PMP12A2h1 gave levels of ADA induction similar to that observed by intraperitoneal injection of L-PMP12A2h1, whereas no ADA production was detected for subcutaneous injection of D-PMP12A2h1.

In this work, the author established a synthetic protocol for mirror-image VHHs by a stepwise NCL process using four peptide segments, in which two conserved cysteines and one alanine were used for the ligation sites. The versatile synthetic route was confirmed by two examples to prepare anti-GFP VHH and PMP12A2h1. Using the established synthetic protocol, radiolabeled PMP12A2h1 and its mirror-image protein were also synthesized via site-specific modification of the C-terminal segment with DTPA. The folded synthetic proteins reproduced the structure and biological functions of native proteins, which were confirmed by spectroscopic and biological analyses. In the comparative biodistribution study of $^{111}$In-DTPA-labeled PMP12A2h1 enantiomers in mice, similar distribution profiles were observed between L-VHH and D-VHH, except for the slightly higher accumulation of D-VHH in the stomach. The reasons behind this difference remain unresolved. A large proportion of administered synthetic L-VHH or D-VHH rapidly accumulated in the kidney. Immunogenicity studies in mice demonstrated that D-VHH induced significantly less ADA generation than L-VHH. These findings of the favorable properties of D-VHHs suggest that D-VHH represents a promising scaffold for developing novel protein therapeutics with lower immunogenicity.

## Experimental Section

**General.** All reagents and solvents were purchased from Watanabe Chemical Industries, Ltd. (Hiroshima, Japan), Kokusan Chemical Industries, Ltd. (Kanagawa, Japan), Sigma-Aldrich JAPAN (Tokyo, Japan), Wako Pure Chemical Industries, Ltd. (Osaka, Japan), Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan) or Nakalai Tesque, Inc. (Kyoto, Japan). For analytical high-performance liquid chromatography (HPLC), a Cosmosil 5C18-AR300 column (4.6 × 250 mm, Nacalai Tesque, Inc.) was employed with a linear gradient of $CH_3CN$ containing 0.1% (v/v) TFA at a flow rate of 1 mL/min (25 °C). For analysis of peptides **10**, **21**, **28**, **29** and **30**, a Cosmosil 5C4-AR300 column (4.6 × 150 mm, Nacalai Tesque, Inc.) was employed with a linear gradient of $CH_3CN$ containing 0.1% (v/v) TFA at a flow rate of 1 mL/min (25 °C or 40 °C). The products were detected by UV absorbance at 220 nm. For preparative HPLC, a Cosmosil 5C18-AR300 column (20 × 250 mm, Nacalai Tesque, Inc.) or a Cosmosil 5C4-AR300 column (20 × 150 mm, Nacalai Tesque, Inc.) was employed with a linear gradient of $CH_3CN$ containing 0.1% TFA at a flow rate of 8 mL/min (room temperature). For purification of peptides L-**10** and D-**10**, Cosmosil 5C4-AR300 column (20 × 150 mm, Nacalai Tesque, Inc.) was employed with a linear gradient of $CH_3CN$ containing 0.1% TFA at a flow rate of 12 mL/min (40 °C). All peptides were characterized by ESI-MS (LCMS-2020, Shimadzu) or MALDI-TOF-MS (microflex, Bruker).

**Solid-Phase Peptide Synthesis.** Standard Fmoc-based solid-phase peptide synthesis (Fmoc-SPPS) was performed using an automatic peptide synthesizer (PSSM-8, Shimadzu). The following side chain protected amino acids were employed: Arg(Pbf), Asn(Trt), Asp(O*t*-Bu), Cys(Trt), Gln(Trt), Glu(O*t*-Bu), His(Trt), Lys(Boc), Ser(*t*-Bu), Thr(*t*-Bu), and Tyr(*t*-Bu). Fmoc-protected amino acids (5 equiv) were coupled using HBTU (5 equiv), $HOBt \cdot H_2O$ (5 equiv) and $(i\text{-}Pr)_2NEt$ (10 equiv) in DMF for 45 min twice. The Fmoc protection group was deprotected by 20% piperidine in DMF for 4 min twice.

**Microwave-Assisted Solid-Phase Peptide Synthesis.** Microwave (MW)-assisted peptide synthesis was performed on a Liberty BLUE (CEM Japan) peptide synthesizer. The following side chain protected amino acids were employed: Arg(Pbf), Asn(Trt), Asp(O*t*-Bu), Cys(Trt), Gln(Trt), Glu(O*t*-Bu), His(Trt), Lys(Boc), Ser(*t*-Bu), Thr(*t*-Bu), Trp(Boc) and Tyr(*t*-Bu). The peptide chain was elongated using the general protocols.[29] Briefly, Fmoc-protected amino acids (5 equiv) except for Arg(Pbf) and His(Trt) were coupled using DIC (5 equiv) and Oxyma pure (5 equiv) in DMF at 90 °C for 2 min. The Fmoc protection group was deprotected by 20% piperidine in DMF for 1 min.

**L-Anti-GFP VHH$^{1–21}$ (L-1).** By the standard procedure of standard Fmoc-SPPS using an automatic peptide synthesizer (PSSM-8, Shimadzu), the peptide sequence was constructed from Rink amide resin (45 mg, 0.020 mmol). After chain assembly, N-terminal amine was protected by $Boc_2O$ (22 mg, 0.10 mmol), (*i*-Pr)$_2$NEt (35 µL, 0.20 mmol) in DMF for 2 h. The resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (2 mL) for 1 h followed by 0.5 M (*i*-Pr)$_2$NEt in DMF (2 mL) for 15 min. Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of MESNa buffer (500 mM MESNa, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 6.0) and incubated for 1 h. The crude products were purified by preparative HPLC to afford the desired peptide L-**1** (9.1 mg, 21% yield). MS (ESI): calcd for $C_{92}H_{159}N_{27}O_{32}S_2$: 2219.56; observed: $[M+3H]^{3+}$ *m/z* = 740.70, $[M+2H]^{2+}$ *m/z* = 1110.55.

**[Cys(Acm)$^{22}$]-L-Anti-GFP VHH$^{22–53}$ (L-4).** By the standard procedure of standard Fmoc-SPPS using an automatic peptide synthesizer (PSSM-8, Shimadzu), the protected peptide was constructed from Rink amide resin (45 mg, 0.020 mmol). For the coupling of an N-terminal cysteine in peptide L-**4**, Boc-Cys(Acm)-OH was employed with HBTU (38 mg, 0.10 mmol), HOBt·$H_2O$ (15 mg, 0.10 mmol), and (*i*-Pr)$_2$NEt (35 µL, 0.20 mmol) in DMF for 2 h. Then, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (2 mL) for 1 h followed by 0.5 M (*i*-Pr)$_2$NEt in DMF (2 mL) for 15 min. Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of MESNa buffer (500 mM MESNa, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 6.0) and incubated for 1 h. The crude products were purified by preparative HPLC to afford desired peptide L-**4** (9.3 mg, 12% yield). MS (ESI): calcd for $C_{166}H_{250}N_{50}O_{48}S_5$: 3874.43; observed: $[M+4H]^{4+}$ *m/z* = 969.40, $[M+3H]^{3+}$ *m/z* = 1292.15, $[M+2H]^{2+}$ *m/z* = 1937.95.

**[Cys$^{54}$]-L-Anti-GFP VHH$^{54–95}$ (L-5).** By the standard procedure of standard Fmoc-SPPS using an automatic peptide synthesizer (PSSM-8, Shimadzu), the peptide sequence was constructed from Rink amide-PEG resin (87 mg, 0.020 mmol). Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum

amount of 50% $CH_3CN$ containing 0.1% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**5** (11 mg, 10% yield). MS (ESI): calcd for $C_{220}H_{340}N_{66}O_{73}S_2$: 5141.65; observed: $[M+6H]^{6+}$ *m/z* = 857.70, $[M+5H]^{5+}$ *m/z* = 1029.05, $[M+4H]^{4+}$ *m/z* = 1286.15, $[M+3H]^{3+}$ *m/z* = 1714.60.

**L-Anti-GFP VHH$^{96–122}$ (L-3).** By the standard procedure of standard Fmoc-SPPS using an automatic peptide synthesizer (PSSM-8, Shimadzu), the peptide sequence was constructed from Rink amide resin (93 mg, 0.040 mmol). Global deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.1% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**3** (3.6 mg, 2.8% yield). MS (ESI): calcd for $C_{137}H_{194}N_{46}O_{38}S$: 3125.40; observed: $[M+4H]^{4+}$ *m/z* = 782.20, $[M+3H]^{3+}$ *m/z* = 1042.55, $[M+2H]^{2+}$ *m/z* = 1563.40.

**[Cys(Acm)$^{22}$/Cys$^{54}$]-L-Anti-GFP VHH$^{22–95}$ (L-6).** Peptide thioester L-**4** (7.7 mg, 2.0 μmol) and peptide L-**5** (11 mg, 2.1 μmol) were reacted in ligation buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 1.0 mL) for 2 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford desired peptide L-**6** (12 mg, 70% yield). MS (ESI): calcd for $C_{384}H_{584}N_{116}O_{118}S_5$: 8873.89; observed: $[M+9H]^{9+}$ *m/z* = 986.75, $[M+8H]^{8+}$ *m/z* = 1110.00, $[M+7H]^{7+}$ *m/z* = 1268.40, $[M+6H]^{6+}$ *m/z* = 1479.75, $[M+5H]^{5+}$ *m/z* = 1775.55.

**Desulfurization of Cys SH group: Synthesis of [Cys(Acm)$^{22}$]-L-Anti-GFP VHH$^{22–95}$ (L-7).** Lyophilized peptide L-**6** (13 mg, 1.5 μmol) was reacted in desulfurization buffer (20 mM VA-044, 100 mM MESNa, 250 mM TCEP, 6 M guanidine·HCl, 100 mM phosphate buffer, pH 6.5; 3.0 mL) for 2 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford desired peptide L-**7** (11 mg, 85% yield). MS (ESI): calcd for $C_{384}H_{584}N_{116}O_{118}S_4$: 8841.83; observed: $[M+9H]^{9+}$ *m/z* = 983.25, $[M+8H]^{8+}$ *m/z* = 1106.00, $[M+7H]^{7+}$ *m/z* = 1263.90, $[M+6H]^{6+}$ *m/z* = 1474.40, $[M+5H]^{5+}$ *m/z* = 1769.10.

**Deprotection of Acm group: Synthesis of L-Anti-GFP VHH$^{22–95}$ (L-2).** Peptide L-**7** (6.9 mg, 0.78 μmol) and $PdCl_2$ (1.4 mg, 7.8 μmol) were reacted in 100 mM phosphate buffer containing 6 M guanidine·HCl (0.39 mL) for 30 min at 37 °C and then added small amounts of DTT. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford desired peptide L-**2** (3.9 mg, 58% yield). MS

(ESI): calcd for $C_{381}H_{579}N_{115}O_{117}S_4$: 8770.75; observed: $[M+9H]^{9+}$ *m/z* = 975.35, $[M+8H]^{8+}$ *m/z* = 1097.15, $[M+7H]^{7+}$ *m/z* = 1253.75, $[M+6H]^{6+}$ *m/z* = 1462.60, $[M+5H]^{5+}$ *m/z* = 1755.00.

**L-Anti-GFP VHH$^{1–95}$ (L-8).** Thioester L-**1** (1.8 mg, 0.80 μmol) and cysteine-peptide L-**2** (2.3 mg, 0.27 μmol) were reacted in ligation buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 0.14 mL) for 1 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide L-**8** (1.8 mg, 62% yield). MS (ESI): calcd for $C_{471}H_{732}N_{142}O_{146}S_4$: 10848.13; observed: $[M+10H]^{10+}$ *m/z* = 1085.60, $[M+9H]^{9+}$ *m/z* = 1206.10, $[M+8H]^{8+}$ *m/z* = 1356.85, $[M+7H]^{7+}$ *m/z* = 1550.45, $[M+6H]^{6+}$ *m/z* = 1808.85.

**L-Anti-GFP VHH$^{1–95}$-MPAA (L-9).** An $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 12.5 μL) was added to a solution of Dbz-peptide L-**8** (1.1 mg, 0.10 μmol) in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 25 μL) at –20 °C, the reaction was continued for 30 min. An MPAA buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 13 μL) was then added to the reaction, and the reaction was continued for 1 min. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide L-**9** (0.66 mg, 61% yield). MS (ESI): calcd for $C_{466}H_{719}N_{135}O_{146}S_5$: 10708.98; observed: $[M+10H]^{10+}$ *m/z* = 1071.65, $[M+9H]^{9+}$ *m/z* = 1190.50, $[M+8H]^{8+}$ *m/z* = 1339.45, $[M+7H]^{7+}$ *m/z* = 1530.65, $[M+6H]^{6+}$ *m/z* = 1785.25.

**L-Anti-GFP VHH (L-10).** Thioester L-**9** (7.6 mg, 0.71 μmol) and cysteine-peptide L-**3** (5.2 mg, 1.7 μmol) were reacted in ligation buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 1.1 mL) for 10 h at room temperature. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide L-**10** (6.0 mg, 62% yield). MS (ESI): calcd for $C_{595}H_{905}N_{181}O_{182}S_5$: 13666.17; observed: $[M+11H]^{11+}$ *m/z* = 1243.20, $[M+10H]^{10+}$ *m/z* = 1367.40, $[M+9H]^{9+}$ *m/z* = 1519.25, $[M+8H]^{8+}$ *m/z* = 1709.95, $[M+7H]^{7+}$ *m/z* = 1952.95.

**Folding of Synthetic Anti-GFP VHH.** Lyophilized peptide L-**10** or D-**10** was dissolved in PBS (pH 7.4) containing 6 M guanidine·HCl and 40 mM DTT at the concentration of 1.0 mg/mL, and the mixture was incubated for 2 h at room temperature. The solution was diluted 100-fold in PBS (pH 7.4) and placed overnight at room temperature. The solution was concentrated using an MWCO 3000 centrifugal filtration membrane (Millipore, Amicon-Ultra 3 kDa).

**PMP12A2h1$^{1–21}$ (L-11).** By the standard procedure of MW-assisted Fmoc-SPPS, the peptide sequence was constructed from Rink amide resin (440 mg, 0.25 mmol). After chain assembly, N-terminal amine was protected by $Boc_2O$ (270 mg, 1.25 mmol), (*i*-$Pr)_2NEt$ (440 µL, 2.5 mmol) in DMF for 2 h. After wash, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (25 mL) for 1 h followed by 0.5 M (*i*-$Pr)_2NEt$ in DMF (25 mL) for 15 min. Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.1% TFA. The crude products were purified by preparative HPLC to afford the desired peptide L-**11** (110 mg, 19% yield). MS (ESI): calcd for $C_{103}H_{169}N_{33}O_{33}$: 2397.68; observed: $[M+4H]^{4+}$ *m/z* = 600.35, $[M+3H]^{3+}$ *m/z* = 800.00, $[M+2H]^{2+}$ *m/z* = 1199.50.

**[Cys(Acm)$^{22}$]-L-PMP12A2h1$^{22–50}$ (L-14).** By the standard procedure of MW-assisted Fmoc-SPPS, the protected peptide was constructed from Rink amide resin (440 mg, 0.25 mmol). For the coupling of an N-terminal cysteine in peptide L-**4**, Boc-Cys(Acm)-OH was employed with DIC/Oxyma in DMF for 2 h. Then, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (25 mL) for 1 h followed by 0.5 M (*i*-$Pr)_2NEt$ in DMF (25 mL) for 15 min. Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.1% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**14** (260 mg, 29% yield). MS (ESI): calcd for $C_{153}H_{229}N_{49}O_{40}S_2$: 3458.94; observed: $[M+5H]^{5+}$ *m/z* = 692.65, $[M+4H]^{4+}$ *m/z* = 865.50, $[M+3H]^{3+}$ *m/z* = 1153.75.

**[Cys$^{50}$]-L-PMP12A2h1$^{50–95}$ (L-15).** By the standard procedure of standard Fmoc-SPPS described above, the peptide sequence was constructed from Rink amide-PEG resin (91 mg, 0.020 mmol). Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.1% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**15** (3.1 mg, 2.8% yield). MS (ESI): calcd for $C_{241}H_{374}N_{72}O_{75}S_3$: 5576.25; observed: $[M+7H]^{7+}$ *m/z* = 797.35, $[M+6H]^{6+}$ *m/z* = 930.05, $[M+5H]^{5+}$ *m/z* = 1116.00, $[M+4H]^{4+}$ *m/z* = 1394.70, $[M+3H]^{3+}$ *m/z* = 1859.55.

**L-PMP12A2h1$^{96–128}$ (L-13).** By the standard procedure of standard Fmoc-SPPS described above, the peptide sequence was constructed from Rink amide-PEG resin (91 mg, 0.020 mmol). Global deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.1% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**13** (6.1 mg, 8.7% yield). MS (ESI): calcd for $C_{150}H_{236}N_{46}O_{49}S$: 3499.87; observed: $[M+3H]^{3+}$ *m/z* = 1167.25, $[M+2H]^{2+}$ *m/z* = 1750.40.

**[Cys(Acm)$^{22}$/Cys$^{50}$]-L-PMP12A2h1$^{22–95}$ (L-16).** MeNbz-peptide L-**14** (17 mg, 4.8 μmol) and cysteine-peptide L-**15** (30 mg, 5.3 μmol) were reacted in ligation buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 2.4 mL) for 2 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford desired peptide L-**16** (12 mg, 30% yield). MS (ESI): calcd for $C_{379}H_{582}N_{114}O_{112}S_5$: 8687.81; observed: $[M+12H]^{12+}$ *m/z* = 724.85, $[M+9H]^{9+}$ *m/z* = 966.05, $[M+8H]^{8+}$ *m/z* = 1086.75, $[M+7H]^{7+}$ *m/z* = 1241.80, $[M+6H]^{6+}$ *m/z* = 1448.70, $[M+5H]^{5+}$ *m/z* = 1738.30.

**[Cys(Acm)$^{22}$]-L-PMP12A2h1$^{22–95}$ (L-17).** By the identical procedure described for the synthesis of L-**7**, lyophilized peptide L-**16** (0.47 mg, 0.054 μmol) was reacted in desulfurization buffer (20 mM VA-044, 100 mM MESNa, 250 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 6.5; 107 μL) for 2 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford desired peptide L-**17** (0.21 mg, 45% yield). MS (ESI): calcd for $C_{379}H_{582}N_{114}O_{112}S_4$: 8655.75; observed: $[M+9H]^{9+}$ *m/z* = 962.50, $[M+8H]^{8+}$ *m/z* = 1082.70, $[M+7H]^{7+}$ *m/z* = 1237.25, $[M+6H]^{6+}$ *m/z* = 1443.35, $[M+5H]^{5+}$ *m/z* = 1731.85.

**L-PMP12A2h1$^{22–95}$ (L-12).** By the identical procedure described for the synthesis of L-**2**, peptide L-**17** (61 mg, 7.0 μmol) and $PdCl_2$ (12 mg, 70 μmol) were reacted in 200 mM phosphate buffer containing 6 M guanidine·HCl (3.5 mL) for 30 min at 37 °C and then added small amounts of DTT. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford desired peptide L-**12** (29 mg, 47% yield). MS (ESI): calcd for $C_{376}H_{577}N_{113}O_{111}S_4$: 8584.67; observed: $[M+9H]^{9+}$ *m/z* = 954.55, $[M+8H]^{8+}$ *m/z* = 1073.80, $[M+7H]^{7+}$ *m/z* = 1227.00, $[M+6H]^{6+}$ *m/z* = 1431.95, $[M+5H]^{5+}$ *m/z* = 1717.55.

**L-PMP12A2h1$^{1–95}$ (L-18).** Nbz-peptide L-**11** (22 mg, 9.3 μmol) and cysteine-peptide L-**12** (27 mg, 3.1 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 1.5 mL) for 1 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide L-**18** (21 mg, 64% yield). MS (ESI): calcd for $C_{465}H_{727}N_{139}O_{141}S_4$: 10649.00; observed: $[M+10H]^{10+}$ *m/z* = 1065.65, $[M+9H]^{9+}$ *m/z* = 1183.95, $[M+8H]^{8+}$ *m/z* = 1331.85, $[M+7H]^{7+}$ *m/z* = 1522.00, $[M+6H]^{6+}$ *m/z* = 1775.60.

**L-PMP12A2h1$^{1–95}$-MPAA (L-19).** By the identical procedure described for the synthesis of L-**9**, an $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 26 μL) was added to a solution of Dbz-peptide L-**18** (2.2 mg, 0.21 μmol) in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 52 μL) at –20 °C, the reaction was continued for 30 min. An MPAA buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 26 μL) was then added to the reaction, and the reaction was continued for 1 min. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide L-**19** (1.1 mg, 52% yield). MS (ESI): calcd for $C_{460}H_{714}N_{132}O_{141}S_5$: 10509.86; observed: $[M+10H]^{10+}$ *m/z* = 1051.75, $[M+9H]^{9+}$ *m/z* = 1168.45, $[M+8H]^{8+}$ *m/z* = 1314.45, $[M+7H]^{7+}$ *m/z* = 1502.20, $[M+6H]^{6+}$ *m/z* = 1752.40.

**L-PMP12A2h1 (SH) (L-20).** Thioester L-**19** (1.1 mg, 0.11 μmol) and cysteine-peptide L-**13** (0.84 mg, 0.24 μmol) were reacted in ligation buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 55 μL) for 2 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide L-**20** (1.0 mg, 68% yield). MS (ESI): calcd for $C_{602}H_{942}N_{178}O_{188}S_5$: 13841.52; observed: $[M+11H]^{11+}$ *m/z* = 1259.05, $[M+10H]^{10+}$ *m/z* = 1384.90, $[M+9H]^{9+}$ *m/z* = 1538.65, $[M+8H]^{8+}$ *m/z* = 1730.95, $[M+7H]^{7+}$ *m/z* = 1978.10.

**Folding of PMP12A2h1.** Lyophilized peptide L-**20** or D-**20** was dissolved in PBS (pH 7.4) containing 6 M guanidine·HCl and 40 mM DTT at the concentration of 1.0 mg/mL, and the mixture was incubated for 2 h at room temperature. The solution was diluted 100-fold in PBS (pH 7.4) and placed overnight at room temperature. The solution was concentrated using an MWCO 3000 centrifugal filtration membrane (Millipore, Amicon-Ultra 3 kDa).

**L-PMP12A2h1 (L-21).** About ten molar excess of 10 mM DTNB in EtOH (10 μL) was added to a solution of folded L-**20** in PBS (30–40 μM, 600–800 μL), and the mixture was incubated for 5 h at 37 °C. The reaction was monitored by analytical HPLC. The crude

products were purified using an MWCO 3000 centrifugal filtration membrane (Millipore, Amicon-Ultra 3 kDa). MS (MALDI-TOF) *m/z* calcd for $C_{602}H_{940}N_{178}O_{188}S_5$ ($[M+H]^+$): 13840.51, found 13840.60.

**[Thz$^{50}$]-L-PMP12A2h1$^{50–74}$ (L-22).** Thioester L-**22** was synthesized via side-chain-anchoring strategy.[40] Fmoc-Asp-OAllyl (1.7 g, 4.4 mmol) was manually loaded onto Rink amide resin (1.5 g, 0.87 mmol) by DIC/Oxyma activation. The peptide sequence was constructed by standard Fmoc-SPPS. Boc protection at N-terminal amine was conducted by treatment with $(Boc)_2O$ (949 mg, 4.4 mmol) and $(i\text{-Pr})_2NEt$ (1.5 mL, 8.7 mmol) at room temperature for 2 h. After washing the resin with $CH_2Cl_2$, DMF and dry $CH_2Cl_2$, allyl protecting group was removed by addition of a solution of $Pd(PPh_3)_4$ (890 mg, 0.77 mmol) and $PhSiH_3$ (4.3 mL, 35 mmol) in dry $CH_2Cl_2$ (17 mL). The resin was shaken for 1 h and the procedure was repeated. Afterwards, the resin was washed with $CH_2Cl_2$, DMF and $CH_2Cl_2$. A solution of ethyl 3-mercaptopropionate (2.6 mL, 21 mmol), HOAt (3.6 g, 26 mmol), $(i\text{-Pr})_2NEt$ (5.7 mL, 33 mmol) and DIC (4.1 mL, 26 mmol) in $CH_2Cl_2$/DMF (6.7 mL, 4:1 by volume) was added and the resin was shaken for 1 h. This thioesterfication step was repeated. Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.1% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**22** (393 mg, 16% yield). MS (ESI): calcd for $C_{123}H_{189}N_{35}O_{43}S_2$: 2910.19; observed: $[M+3H]^{3+}$ *m/z* = 970.75, $[M+2H]^{2+}$ *m/z* = 1455.75.

**[Cys$^{75}$]-L-PMP12A2h1$^{75–95}$ (L-23).** By the identical procedure described for the synthesis of L-**15**, peptide L-**23** was synthesized (110 mg, 16% yield) from Rink amide resin (440 mg, 0.25 mmol). MS (ESI): calcd for $C_{124}H_{195}N_{37}O_{34}S_3$: 2844.33; observed: $[M+4H]^{4+}$ *m/z* = 711.90, $[M+3H]^{3+}$ *m/z* = 948.80, $[M+2H]^{2+}$ *m/z* = 1422.80.

**[Cys$^{50}$/Cys$^{75}$]-L-PMP12A2h1$^{50–95}$ (L-24).** Thioester L-**22** (140 mg, 48 μmol) and cysteine-peptide L-**23** (130 mg, 44 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 22 mL) for 9 h at room temperature. Then, a $MeONH_2$ solution (1 M $MeONH_2$, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 4.0; 22 mL) was added, and reaction mixture was incubated for 9 h at room temperature. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford desired peptide L-**24** (89 mg, 36% yield). MS (ESI): calcd for $C_{241}H_{374}N_{72}O_{75}S_4$: 5608.31;

observed: $[M+7H]^{7+}$ *m/z* = 801.85, $[M+5H]^{5+}$ *m/z* = 1122.40, $[M+4H]^{4+}$ *m/z* = 1402.75, $[M+3H]^{3+}$ *m/z* = 1870.30.

**[Cys(Acm)$^{22}$/Cys$^{50}$/Cys$^{75}$]-L-PMP12A2h1$^{22–95}$ (L-25).** MeNbz-peptide L-**14** (50 mg, 15 μmol) and cysteine-peptide L-**24** (89 mg, 16 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 2.4 mL) for 2 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford desired peptide L-**25** (79 mg, 62% yield). MS (ESI): calcd for $C_{379}H_{582}N_{114}O_{112}S_6$: 8719.87; observed: $[M+9H]^{9+}$ *m/z* = 969.70, $[M+8H]^{8+}$ *m/z* = 1090.75, $[M+7H]^{7+}$ *m/z* = 1246.30, $[M+6H]^{6+}$ *m/z* = 1454.25, $[M+5H]^{5+}$ *m/z* = 1744.80.

**DTPA-L-PMP12A2h1$^{96–128}$ (L-27).** Fmoc-Lys(Mtt)-OH (62 mg, 0.10 mmol) was manually loaded onto NovaSyn TGR resin (80 mg, 0.020 mmol) by HATU (38 mg, 0.10 mmol) and (*i*-Pr)$_2$NEt (34 μL, 0.20 mmol) in DMF for 2 h. Then the Mtt group was deprotected by treatment of $CH_2Cl_2$/TFA/TIS (94:1:5) at room temperature for 5 min × 8. *N,N,N'',N''*-Tetrakis[(*tert*-butyloxycarbonyl)methyl]-*N'*-[acetic acid]diethylenetriamine (**26**)[37] (25 mg, 0.040 mmol) was coupled by HATU (15mg, 0.040 mmol) and (*i*-Pr)$_2$NEt (14 μL, 0.080 mmol) in DMF at 37 °C for 2 h twice. After Fmoc removal of Lys, Fmoc-Ahx-OH (35 mg, 0.10 mmol) was then coupled by the identical HATU/(*i*-Pr)$_2$NEt activation. The peptide sequence was constructed by standard Fmoc-SPPS. Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.1% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**27** (6.0 mg, 7.1% yield). MS (ESI): calcd for $C_{182}H_{291}N_{53}O_{61}S$: 4229.70; observed: $[M+5H]^{5+}$ *m/z* = 846.75, $[M+4H]^{4+}$ *m/z* = 1058.15, $[M+3H]^{3+}$ *m/z* = 1410.55.

**DTPA-L-PMP12A2h1 (SH form) (L-28).** By the identical procedure described for the synthesis of L-**20**, thioester L-**19** (1.7 mg, 0.16 μmol) and cysteine-peptide L-**27** (0.94 mg, 0.23 μmol) were converted into L-**28** (1.0 mg, 43% yield). MS (ESI): calcd for $C_{634}H_{997}N_{185}O_{200}S_5$: 14571.35; observed: $[M+12H]^{12+}$ *m/z* = 1215.15, $[M+11H]^{11+}$ *m/z* = 1325.35, $[M+10H]^{10+}$ *m/z* = 1457.90, $[M+9H]^{9+}$ *m/z* = 1619.75, $[M+8H]^{8+}$ *m/z* = 1821.90.

**DTPA-L-PMP12A2h1 (SS form) (L-29).** 10 mM DTNB in EtOH (36 μL) was added to a solution of L-**28** (1.0 mg, 71 nmol) in PBS (pH 7.4) containing 6 M guanidine·HCl (10 mL), and the mixture was incubated for 10 min at room temperature. The reaction was

monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide L-**29** (0.51 mg, 49% yield). MS (MALDI-TOF) *m/z* calcd for $C_{634}H_{995}N_{185}O_{200}S_5$ ($[M+H]^+$): 14570.34, found 14570.60.

**Folding of DTPA-PMP12A2h1.** Lyophilized peptide L-**29** or D-**29** was dissolved in PBS (pH 7.4) containing 6 M guanidine·HCl at the concentration of 1.0 mg/mL, and the mixture was incubated for 2 h at room temperature. The solution was diluted 100-fold in PBS (pH 7.4) and placed overnight at room temperature. The solution was concentrated using an MWCO 3000 centrifugal filtration membrane (Millipore, Amicon-Ultra 3 kDa).

**[In]-DTPA-L-PMP12A2h1 (L-30).** Folded L-**29** was incubated with a 5-fold molar excess of $InCl_3$ in 0.1 M $NH_4OAc$ (pH 5.0) for 2 h at room temperature. The desired In-labeled L-**30** was purified using an MWCO 3000 centrifugal filtration membrane (Millipore, Amicon-Ultra 10 kDa). MS (MALDI-TOF) *m/z* calcd for $C_{634}H_{995}InN_{185}O_{200}S_5$ ($[M+H]^+$): 14685.15, found 14685.30.

**D-Anti-GFP VHH$^{1–21}$ (D-1).** By the identical procedure described for the synthesis of L-**1**, peptide D-**1** was synthesized (60.9 mg, 20% yield) from Rink amide resin (310 mg, 0.14 mmol). MS (ESI): calcd for $C_{92}H_{159}N_{27}O_{32}S_2$: 2219.56; observed: $[M+3H]^{3+}$ *m/z* = 740.75, $[M+2H]^{2+}$ *m/z* = 1110.40.

**[D-Cys(Acm)$^{22}$]-D-Anti-GFP VHH$^{22–53}$ (D-4).** By the identical procedure described for the synthesis of L-**4**, peptide D-**4** was synthesized (68.6 mg, 13% yield) from Rink amide resin (360 mg, 0.16 mmol). MS (ESI): calcd for $C_{166}H_{250}N_{50}O_{48}S_5$: 3874.43; observed: $[M+4H]^{4+}$ *m/z* = 969.35, $[M+3H]^{3+}$ *m/z* = 1292.15, $[M+2H]^{2+}$ *m/z* = 1938.00.

**[D-Cys$^{54}$]-D-Anti-GFP VHH$^{54–95}$ (D-5).** By the identical procedure described for the synthesis of L-**5**, peptide D-**5** was synthesized (7.2 mg, 7.0% yield) from Rink amide-PEG resin (87 mg, 0.020 mmol). MS (ESI): calcd for $C_{220}H_{340}N_{66}O_{73}S_2$: 5141.65; observed: $[M+7H]^{7+}$ *m/z* = 735.25, $[M+6H]^{6+}$ *m/z* = 857.70, $[M+5H]^{5+}$ *m/z* = 1029.05, $[M+4H]^{4+}$ *m/z* = 1286.10, $[M+3H]^{3+}$ *m/z* = 1714.55.

**D-Anti-GFP VHH$^{96–122}$ (D-3).** By the identical procedure described for the synthesis of L-**3**, peptide D-**3** was synthesized (10.2 mg, 2.3% yield) from Rink amide-PEG resin (610 mg, 0.14 mmol). MS (ESI): calcd for $C_{137}H_{194}N_{46}O_{38}S$: 3125.40; observed: $[M+4H]^{4+}$ *m/z* = 782.10, $[M+3H]^{3+}$ *m/z* = 1042.55, $[M+2H]^{2+}$ *m/z* = 1563.40.

**[D-Cys(Acm)$^{22}$/D-Cys$^{54}$]-D-Anti-GFP VHH$^{22–95}$ (D-6).** By the identical procedure described for the synthesis of L-**6**, peptide D-**4** (9.1 mg, 2.4 μmol) and peptide D-**5** (12 mg, 2.4 μmol) were converted into D-**6** (13.8 mg, 65% yield). MS (ESI): calcd for

$C_{384}H_{584}N_{116}O_{118}S_5$: 8873.89; observed: $[M+9H]^{9+}$ *m/z* = 986.70, $[M+8H]^{8+}$ *m/z* = 1109.95, $[M+7H]^{7+}$ *m/z* = 1268.45, $[M+6H]^{6+}$ *m/z* = 1479.65, $[M+5H]^{5+}$ *m/z* = 1775.55.

**[D-Cys(Acm)$^{22}$]-D-Anti-GFP VHH$^{22–95}$ (D-7).** By the identical procedure described for the synthesis of L-**7**, peptide D-**6** (8.0 mg, 0.90 μmol) was converted into D-**7** (3.7 mg, 47% yield). MS (ESI): calcd for $C_{384}H_{584}N_{116}O_{118}S_4$: 8841.83; observed: $[M+9H]^{9+}$ *m/z* = 983.15, $[M+8H]^{8+}$ *m/z* = 1106.00, $[M+7H]^{7+}$ *m/z* = 1263.90, $[M+6H]^{6+}$ *m/z* = 1474.40, $[M+5H]^{5+}$ *m/z* = 1769.10.

**D-Anti-GFP VHH$^{22–95}$ (D-2).** By the identical procedure described for the synthesis of L-**2**, peptide D-**7** (5.0 mg, 0.56 μmol) was converted into D-**2** (3.1 mg, 63% yield). MS (ESI): calcd for $C_{381}H_{579}N_{115}O_{117}S_4$: 8770.75; observed: $[M+11H]^{11+}$ *m/z* = 798.15, $[M+10H]^{10+}$ *m/z* = 877.90, $[M+9H]^{9+}$ *m/z* = 975.35, $[M+8H]^{8+}$ *m/z* = 1097.05, $[M+7H]^{7+}$ *m/z* = 1253.75.

**D-Anti-GFP VHH$^{1–95}$ (D-8).** By the identical procedure described for the synthesis of L-**8**, peptide D-**1** (3.0 mg, 1.4 μmol) and peptide D-**2** (3.9 mg, 0.44 μmol) were converted into D-**8** (3.6 mg, 75% yield). MS (ESI): calcd for $C_{471}H_{732}N_{142}O_{146}S_4$: 10848.13; observed: $[M+10H]^{10+}$ *m/z* = 1085.40, $[M+9H]^{9+}$ *m/z* = 1206.15, $[M+8H]^{8+}$ *m/z* = 1356.65, $[M+7H]^{7+}$ *m/z* = 1550.35, $[M+6H]^{6+}$ *m/z* = 1808.60.

**D-Anti-GFP VHH$^{1–95}$-MPAA (D-9).** By the identical procedure described for the synthesis of L-**9**, peptide D-**8** (11 mg, 1.0 μmol) was converted into D-**9** (7.1 mg, 66% yield). MS (ESI): calcd for $C_{466}H_{719}N_{135}O_{146}S_5$: 10708.98; observed: $[M+10H]^{10+}$ *m/z* = 1071.70, $[M+9H]^{9+}$ *m/z* = 1190.55, $[M+8H]^{8+}$ *m/z* = 1339.30, $[M+7H]^{7+}$ *m/z* = 1530.45, $[M+6H]^{6+}$ *m/z* = 1785.70.

**D-Anti-GFP VHH (D-10).** By the identical procedure described in the synthesis of L-**10**, D-**9** (7.1 mg, 0.66 μmol) and D-**3** (3.1 mg, 0.99 μmol) were converted into D-**10** (3.8 mg, 42% yield). MS (ESI): calcd for $C_{595}H_{905}N_{181}O_{182}S_5$: 13666.17; observed: $[M+11H]^{11+}$ *m/z* = 1243.15, $[M+10H]^{10+}$ *m/z* = 1367.45, $[M+9H]^{9+}$ *m/z* = 1519.30, $[M+8H]^{8+}$ *m/z* = 1709.35, $[M+7H]^{7+}$ *m/z* = 1952.85.

**D-PMP12A2h1$^{1–21}$ (D-11).** By the identical procedure described for the synthesis of L-**11**, peptide D-**11** was synthesized (115 mg, 19% yield) from Rink amide resin (439 mg, 0.25 mmol). MS (ESI): calcd for $C_{103}H_{169}N_{33}O_{33}$: 2397.68; observed: $[M+4H]^{4+}$ *m/z* = 600.25, $[M+3H]^{3+}$ *m/z* = 800.00, $[M+2H]^{2+}$ *m/z* = 1199.60.

**[D-Cys(Acm)$^{22}$]-D-PMP12A2h1$^{22–50}$ (D-14).** By the identical procedure described for the synthesis of L-**14**, peptide D-**14** was synthesized (268 mg, 31% yield) from Rink amide

resin (440 mg, 0.25 mmol). MS (ESI): calcd for $C_{153}H_{229}N_{49}O_{40}S_2$: 3458.94; observed: $[M+5H]^{5+}$ *m/z* = 692.60, $[M+4H]^{4+}$ *m/z* = 865.45, $[M+3H]^{3+}$ *m/z* = 1153.65.

**[D-Cys$^{50}$]-D-PMP12A2h1$^{50–95}$ (D-15).** By the identical procedure described for the synthesis of L-**15**, peptide D-**15** was synthesized (307 mg, 5.5% yield) from NovaSyn TGR resin (4.0 g, 1.0 mmol). MS (ESI): calcd for $C_{241}H_{374}N_{72}O_{75}S_3$: 5576.25; observed: $[M+6H]^{6+}$ *m/z* = 930.05, $[M+5H]^{5+}$ *m/z* = 1115.95, $[M+4H]^{4+}$ *m/z* = 1394.80, $[M+3H]^{3+}$ *m/z* = 1859.60.

**D-PMP12A2h1$^{96–128}$ (D-13).** By the identical procedure described for the synthesis of L-**13**, peptide D-**13** was synthesized (12 mg, 6.1% yield) from NovaSyn TGR resin (220 mg, 0.054 mmol). MS (ESI): calcd for $C_{150}H_{236}N_{46}O_{49}S$: 3499.87; observed: $[M+3H]^{3+}$ *m/z* = 1167.30, $[M+2H]^{2+}$ *m/z* = 1750.55.

**[D-Cys(Acm)$^{22}$/D-Cys$^{50}$]-D-PMP12A2h1$^{22–95}$ (D-16).** By the identical procedure described for the synthesis of L-**16**, peptide D-**14** (124 mg, 36 μmol) and peptide D-**15** (220 mg, 40 μmol) were converted into D-**16** (139 mg, 45% yield). MS (ESI): calcd for $C_{379}H_{582}N_{114}O_{112}S_5$: 8687.81; observed: $[M+9H]^{9+}$ *m/z* = 966.05, $[M+8H]^{8+}$ *m/z* = 1086.80, $[M+7H]^{7+}$ *m/z* = 1241.85, $[M+6H]^{6+}$ *m/z* = 1448.75, $[M+5H]^{5+}$ *m/z* = 1738.20.

**[D-Cys(Acm)$^{22}$]-D-PMP12A2h1$^{22–95}$ (D-17).** By the identical procedure described for the synthesis of L-**17**, peptide D-**16** (95 mg, 11 μmol) was converted into D-**17** (71 mg, 75% yield). MS (ESI): calcd for $C_{379}H_{582}N_{114}O_{112}S_4$: 8655.75; observed: $[M+9H]^{9+}$ *m/z* = 962.45, $[M+8H]^{8+}$ *m/z* = 1082.70, $[M+7H]^{7+}$ *m/z* = 1237.20, $[M+6H]^{6+}$ *m/z* = 1443.25, $[M+5H]^{5+}$ *m/z* = 1731.85.

**D-PMP12A2h1$^{22–95}$ (D-12).** By the identical procedure described for the synthesis of L-**12**, peptide D-**17** (51 mg, 5.9 μmol) was converted into D-**12** (39 mg, 76% yield). MS (ESI): calcd for $C_{376}H_{577}N_{113}O_{111}S_4$: 8584.67; observed: $[M+9H]^{9+}$ *m/z* = 954.60, $[M+8H]^{8+}$ *m/z* = 1073.80, $[M+7H]^{7+}$ *m/z* = 1227.05, $[M+6H]^{6+}$ *m/z* = 1431.45, $[M+5H]^{5+}$ *m/z* = 1717.60.

**D-PMP12A2h1$^{1–95}$ (D-18).** By the identical procedure described for the synthesis of L-**18**, peptide D-**11** (62 mg, 26 μmol) and peptide D-**12** (44 mg, 5.2 μmol) were converted into D-**18** (39 mg, 70% yield). MS (ESI): calcd for $C_{465}H_{727}N_{139}O_{141}S_4$: 10649.00; observed: $[M+10H]^{10+}$ *m/z* = 1065.70, $[M+9H]^{9+}$ *m/z* = 1183.95, $[M+8H]^{8+}$ *m/z* = 1331.90, $[M+7H]^{7+}$ *m/z* = 1522.05, $[M+6H]^{6+}$ *m/z* = 1775.50.

**D-PMP12A2h1$^{1–95}$-MPAA (D-19).** By the identical procedure described for the synthesis of L-**19**, peptide D-**18** (18 mg, 1.7 μmol) was converted into D-**19** (10 mg, 58% yield). MS

(ESI): calcd for $C_{460}H_{714}N_{132}O_{141}S_5$: 10509.86; observed: $[M+10H]^{10+}$ *m/z* = 1051.75, $[M+9H]^{9+}$ *m/z* = 1168.50, $[M+8H]^{8+}$ *m/z* = 1314.40, $[M+7H]^{7+}$ *m/z* = 1502.05, $[M+6H]^{6+}$ *m/z* = 1752.20.

**D-PMP12A2h1 (SH) (D-20).** By the identical procedure described in the synthesis of L-**20**, peptide D-**19** (4.9 mg, 0.47 μmol) and D-**13** (3.6 mg, 1.0 μmol) were converted into D-**20** (2.6 mg, 40% yield). MS (ESI): calcd for $C_{602}H_{942}N_{178}O_{188}S_5$: 13841.52; observed: $[M+11H]^{11+}$ *m/z* = 1259.05, $[M+10H]^{10+}$ *m/z* = 1384.90, $[M+9H]^{9+}$ *m/z* = 1538.65, $[M+8H]^{8+}$ *m/z* = 1731.10, $[M+7H]^{7+}$ *m/z* = 1978.20.

**D-PMP12A2h1 (D-21).** By the identical procedure described for the synthesis of L-**21**, peptide D-**21** was synthesized. MS (MALDI-TOF) *m/z* calcd for $C_{602}H_{940}N_{178}O_{188}S_5$ ($[M+H]^+$): 13840.51, found 13839.93.

**DTPA-D-PMP12A2h1$^{96–128}$ (D-27).** By the identical procedure described for the synthesis of L-**27**, peptide D-**27** was synthesized (7.9 mg, 9.4% yield) from NovaSyn TGR resin (80 mg, 0.020 mmol). MS (ESI): calcd for $C_{182}H_{291}N_{53}O_{61}S$: 4229.70; observed: $[M+5H]^{5+}$ *m/z* = 846.85, $[M+4H]^{4+}$ *m/z* = 1058.15, $[M+3H]^{3+}$ *m/z* = 1410.55.

**DTPA-D-PMP12A2h1 (SH) (D-28).** By the identical procedure described for the synthesis of L-**28**, thioester D-**19** (1.6 mg, 0.15 μmol) and cysteine-peptide D-**27** (1.0 mg, 0.24 μmol) were converted into D-**28** (1.2 mg, 53% yield). MS (ESI): calcd for $C_{634}H_{997}N_{185}O_{200}S_5$: 14571.35; observed: $[M+13H]^{13+}$ *m/z* = 1121.65, $[M+12H]^{12+}$ *m/z* = 1215.00, $[M+11H]^{11+}$ *m/z* = 1325.55, $[M+10H]^{10+}$ *m/z* = 1458.35, $[M+9H]^{9+}$ *m/z* = 1619.75.

**DTPA-D-PMP12A2h1 (D-29).** By the identical procedure described for the synthesis of L-**29**, peptide D-**28** (1.2 mg, 85 nmol) was converted into D-**29** (0.54 mg, 43% yield). MS (MALDI-TOF) *m/z* calcd for $C_{634}H_{995}N_{185}O_{200}S_5$ ($[M+H]^+$): 14570.34, found 14570.70.

**[In]-DTPA-D-PMP12A2h1 (D-30).** By the identical procedure described for the synthesis of L-**30**, peptide D-**30** was synthesized. MS (MALDI-TOF) *m/z* calcd for $C_{634}H_{995}InN_{185}O_{200}S_5$ ($[M+H]^+$): 14685.15, found 14684.82.

**Preparation of Recombinant Proteins.** The recombinant anti-GFP VHH was prepared according to the reported procedure.[41] Briefly, the plasmid of anti-GFP VHH was prepared by cloning 6×His at the C-terminus end into pGEX6P1-GFP-Nanobody (Addgene #61838). All plasmids were transformed into BL21 (DE3) cells and grown in LB media. Bacteria were cultured in a LB medium with ampicillin (final concentration; 50 μg/mL) overnight at 37 °C and the expression was induced with 100 mM IPTG overnight at 20 °C. Bacteria were pelleted by centrifugation (6,000 rpm, 20 min, 4 °C)

and resuspended in binding buffer (5 mM DTT and protease inhibitor [Nakalai Tesque, Inc.] in PBS). After sonication and addition of Triton X-100, the cell lysates were centrifuged at 12,000 rpm for 30 min at 4 ℃. The supernatants were incubated with Glutathione-Sepharose 4B (GE Healthcare) overnight at 4 °C. For purification of recombinant anti-GFP VHH, GST-tags were removed by incubation with PreScission protease (GE Helthcare) in cleavage buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.01% Triton X-100) overnight at 4 °C. For purification of GST-EGFP and GST-mCherry, GST-tagged proteins were eluted by elution buffer (10 mM glutathione in PBS, pH 8.5) for 30 min at 4 °C. GST-mCherry proteins were further purified by gel filtration chromatography (Superdex 75 Increase, GE Healthcare) in PBS. All purities of proteins were confirmed by SDS-PAGE.

**Measurement of CD Spectra.** Anti-GFP VHHs or PMP12A2h1s were diluted in PBS (pH 7.4), and the concentration was adjusted to 10 μM. CD spectra of proteins were recorded using a JASCO J-720 circular dichroism spectrometer (JASCO, Tokyo, Japan) at 20 °C. The thermal stability of each protein was estimated by monitoring the change in ellipticity between 37 and 95 °C at a wavelength of 203 nm.

**Enzyme-linked Immunosorbent Assay (ELISA).** ELISAs were performed in PBS (pH 7.4) containing 0.025% Tween 20 for all wash and dilution processes. 96-well microtiter plates (Greiner, high binding) were coated overnight at 4 ℃ with either GST-EGFP or GST-mCherry in 50 mM sodium carbonate buffer (pH 9.4) (50 μL/well; 30 nM). After coating, wells were washed three times and blocked with PBS containing 3% BSA (150 μL/well) for 2 h. After three times washes, either recombinant or synthetic L-anti-GFP VHH (50 μL/well; 0.03–30 nM) was added and incubated for 1 h. After three times washes, 1:5,000 dilution of anti 6×histidine monoclonal antibody (28-75) (Wako) (50 μL/well) was added and incubated for 1 h at 4 ℃. After three times washes, 1:5,000 dilution of anti-mouse IgG (H+L)-HRP conjugate (Promega) (50 μL/well) was added and incubated for 1 h at 4 °C. After five times washes, TMB solution was added and incubated for 15 min, then the reaction was stopped using 1 M $H_2SO_4$ (50 μL/well). Absorbance was measured at 450 nm using an iMark Microplate Absorbance Reader (Bio-Rad).

**Quartz Crystal Microbalance (QCM) Analysis.** QCM analysis was performed using a Single-Q instrument (SCINICS, Tokyo, Japan). The gold surface of a sensor chip was cleaned with Piranha solution, and 0.2 mM NTA-SAM Forming Reagent (Dojindo, Kumamoto, Japan) in 10% EtOH was added, and the sample was incubated overnight. After washing with 1 mM NaOH, 40 mM $NiSO_4$ was added and incubated for 1 h. After the sample was washed and equilibrated with 500 μL of PBS (pH 7.4), 5 μL of 1.0 mg/mL

ligand solution (recombinant L-anti-GFP VHH, synthetic L-anti-GFP VHH, or D-anti-GFP VHH in PBS) was injected. After washing and equilibrating with 500 µL of PBS (pH 7.4), each concentration of the analyte solution (final concentration: 0.1–300 nM) was injected repeatedly. The dissociation constants were evaluated from triplicate assays.

**Surface Plasmon Resonance (SPR) Analysis.** SPR analysis of the binding of synthetic PMP12A2h1 to the human VWF-A1 domain was carried out using a Biacore X100 SPR instrument (Cytiva, Tokyo, Japan). Recombinant human VWF-A1 domain (residues 1261–1468), VWD type 2B, was purchased from U-Protein Express BV (Utrecht, Netherlands). HBS-EP buffer (Cytiva) was used as the running buffer at 25 °C, and 1 M NaCl solution was used as the regeneration solution. For evaluating PMP12A2h1 proteins, the VWF-A1 domain was immobilized on a CM5 sensor chip (2500 RU) using an amine coupling kit (EDC/NHS). All analytes were evaluated for 3 min of contact time, followed by 7 min of dissociation at a flow rate of 30 µL/min. $K_D$ values were calculated by steady-state binding from triplicate assays.

**Labeling with $^{111}$Indium: Synthesis of [$^{111}$In]-DTPA-L-PMP12A2h1 (L-30) and [$^{111}$In]-DTPA-D-PMP12A2h1 (D-30).** Folded L-**29** and D-**29** (0.50 mg/mL, 400 µL) were radiolabeled with $^{111}$Indium Chloride Injection (Nihon Medi-Physics Co., Ltd., 37 MBq/500 µL) by incubation for 1 h at room temperature. Radiochemical purity was analyzed by radio-HPLC. An LD-20AD (Shimadzu, Kyoto, Japan) was used for high performance liquid chromatography (HPLC), along with a CBM-20A (Shimadzu) communication bus module, DGU-20A3R (Shimadzu) degassing unit, CTO-20AC (Shimadzu) column oven, and SPD-20A (Shimadzu) ultraviolet (UV) detector (k = 254 nm), and g-survey meter TCS-172 (ALOKA, Mitaka, Japan). The radiolabeled tracer was further purified using an MWCO 10,000 centrifugal filtration membrane (Millipore, Amicon-Ultra 10 kDa).

**Biodistribution Study.** Animal care and experimental procedures were approved by the Animal Care Committee of Kyoto Pharmaceutical University. Six-week-old male ddY mice were purchased from Japan SLC, Inc. A solution of [$^{111}$In]-L-**30** or [$^{111}$In]-D-**30** (37 kBq/100 µL) in a saline solution was injected intravenously into the tails of ddY mice (6 weeks old). The mice were euthanized 15 min, 30 min, 1 h, 2 h, 6 h, and 24 h post-injection ($N$ = 5 in each group). The blood, heart, lungs, stomach, small intestine, colon, liver, spleen, pancreas, kidney, bone, muscle, and brain of each animal were harvested and weighed. The radioactivity in the blood and organs was measured using a 1480 Automatic Gamma Counter WizardTM 3 (PerkinElmer, Waltham). The concentration of radioactivity in each organ was expressed as the percentage of the injected dose per gram.

**Immunization and Immunogenicity Assays.** Animal care and experimental procedures were approved by the Animal Care Committee of Kyoto University. Six-week-old female BALB/c mice were purchased from Japan SLC, Inc. (Hamamatsu, Japan). Six BALB/c mice (female, 6 weeks old) per group were given intraperitoneal injections on day 0 with 50 μg/injection of synthetic L-anti-GFP VHH or D-anti-GFP VHH antigen emulsified in Freund's Complete Adjuvant (Wako, Osaka, Japan) and on days 14 and 21 with 50 μg/injection of synthetic L-anti-GFP VHH or D-anti-GFP VHH antigen emulsified in Freund's Incomplete Adjuvant (Wako). Immune sera were collected on days 0, 14, 21, and 28. ELISAs were performed in PBS (pH 7.4) containing 0.025% Tween 20 for all wash and dilution processes. 96-well microtiter plates (Greiner, high binding) were coated overnight at 4 °C with either L-anti-GFP VHH or D-anti-GFP VHH in 50 mM sodium carbonate buffer (pH 9.4, 50 μL/well; 1.0–1,000 ng/mL). After coating, the wells were washed three times and blocked with PBS containing 3% BSA (150 μL/well) for 2 h. After three washes, 1:1,000 dilution of immunized serum from each mouse (50 μL/well; days 0, 14, 21 and 28) was added and incubated for 1 h. After three washes, 1:5,000 dilution of anti-mouse IgG (H + L)-HRP conjugate (Promega, Madison, WI) (50 μL/well) was added and incubated for 30 min. After four washes, a TMB solution was added and incubated for 15 min, and then the reaction was stopped using 1 M $H_2SO_4$ (50 μL/well). Absorbance was measured at 450 nm using an iMark Microplate Absorbance Reader (Bio-Rad, Hercules, CA). The immunogenicity of L-PMP12A2h1 and D-PMP12A2h1 was investigated by an identical procedure with a few modifications. In this PMP12A2h1 experiment, mice were injected intraperitoneally ($N = 5$) or subcutaneously ($N = 5$) on days 0, 14, and 28 with 25 μg/injection of L-PMP12A2h1 or D-PMP12A2h1. Immune sera were collected on days 0, 14, 28, and 35.

References and Footnotes


1. Hamers-Casterman, C.; Atarhouch, T.; Muyldermans, S.; Robinson, G.; Hamers, C.; Songa, E. B.; Bendahman, N.; Hamers, R. *Nature* **1993**, *363*, 446-448.
2. Mitchell, L. S.; Colwell, L. J. *Proteins* **2018**, *86*, 697–706.
3. Mitchell, L. S.; Colwell, L. J. *Protein Eng. Des. Sel.* **2018**, *31*, 267–275.
4. Jin, B.-k.; Odongo, S.; Radwanska, M.; Magez, S. *Int. J. Mol. Sci.* **2023**, *24*, 5994.
5. Duggan, S. *Drugs* **2018**, *78*, 1639–1642.
6. Keam, S. J. *Drugs* **2023**, *83*, 87–92.
7. Rossotti, M. A.; Belanger, K.; Henry, K. A.; Tanha, J. *FEBS J.* **2022**, *289*, 4304–4327.
8. Tanaka, Y.; Kawanishi, M.; Nakanishi, M.; Yamasaki, H.; Takeuchi, T. *Mod. Rheumatol.* **2023**, *33*, 875–882.
9. Liu, B.; Yang, D. *Int. J. Mol. Sci.* **2022**, *23*, 1482.
10. McMahon, C.; Baier, A. S.; Pascolutti, R.; Wegrecki, M.; Zheng, S.; Ong, J. X.; Erlandson, S. C.; Hilger, D.; Rasmussen, S. G. F.; Ring, A. M.; Manglik, A.; Kruse, A. C. *Nat. Struct. Mol. Biol.* **2018**, *25*, 289–296.
11. Doshi, R.; Chen, B. R.; Vibat, C. R. T.; Huang, N.; Lee, C.-W.; Chang, G. *Sci. Rep.* **2014**, *4*, 6760.
12. Rothbauer, U.; Zolghadr, K.; Tillib, S.; Nowak, D.; Schermelleh, L.; Gahl, A.; Backmann, N.; Conrath, K.; Muyldermans, S.; Cardoso, M. C.; Leonhardt, H. *Nat. Methods* **2006**, *3*, 887–889.
13. Kubala, M. H.; Kovtun, O.; Alexandrov, K.; Collins, B. M. *Protein Sci.* **2010**, *19*, 2389–2401.
14. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.
15. Shu, K.; Iwamoto, N.; Honda, K.; Kondoh, Y.; Hirano, H.; Osada, H.; Ohno, H.; Fujii, N.; Oishi, S. *Bioconjugate Chem.* **2019**, *30*, 1395–1404.
16. Blanco-Canosa, J.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2008**, *47*, 6851–6855.
17. Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248–9252.
18. The author chose $Ala^{54}$ and $Ala^{50}$ for the NCL sites in CDR2 for the synthesis of anti-GFP VHH and PMP12A2h1, respectively, to avoid the possibly low efficient NCL between conserved $Val^{49}$ and $Ala^{50}$. For library design for screening, an alanine may be positioned around this region for the NCL site.
19. It was reported that epimerization often occurs during NCL for the formation of Ser-Cys junctions. In the NCLs using Ser thioesters (L-**1**, L-**4** and L-**11**), the epimerized

products were removed by HPLC purification, to see: Snella, B.; Diemer, V.; Drobecq, H.; Agouridas, V.; Melnyk, O. *Org. Lett.* **2018**, *20*, 7616–7619.
20. Maity, S. K.; Jbara, M.; Laps, S.; Brik, A. *Angew. Chem., Int. Ed.* **2016**, *55*, 8108–8112.
21. Wang, J.-X.; Fang, G.-M.; He, Y.; Qu, D.-L.; Yu, M.; Hong, Z.-Y.; Liu. L. *Angew. Chem., Int. Ed.* **2015**, *54*, 2194–2198.
22. Hartmann, L.; Botzanowski, T.; Galibert, M.; Jullian, M.; Chabrol, E.; Zeder-Lutz, G.; Kugler, V.; Stojko, J.; Strub, J.-M.; Ferry, G.; Frankiewicz, L.; Puget, K.; Wagner, R.; Cianférani, S.; Boutin, J. A. *Protein Sci.* **2019**, *28*, 1865–1879.
23. Huppelschoten, Y.; Elhebieshy, A. F.; Hameed, D. S.; Sapmaz, A.; Buchardt, J.; Nielsen, T. E.; Ovaa, H.; van der Heden van Noort, G. J. *ChemBioChem* **2022**, e202200304.
24. Bruce, V. J.; Lopez-Islas, M.; McNaughton, B. R. *Protein Sci.* **2016**, *25*, 1129–1137.
25. QCM analysis was employed to evaluate the binding affinity of anti-GFP VHH, which was immobilized on the gold surface using the His-tag at the C-terminus.
26. In this model study, the author synthesized mirror-image proteins of anti-GFP VHH and PMP12A2h1 (D-anti-GFP VHH and D-PMP12A2h1). Because the sequence was not derived from mirror-image screening using D-EGFP and D-vWF-A1 domain, these D-VHHs showed no binding to native proteins.
27. Ulrichts, H.; Silence, K.; Schoolmeester, A.; de Jaegere, P.; Rossenu, S.; Roodt, J.; Priem, S.; Lauwereys, M.; Casteels, P.; Bockstaele, F. V.; Verschueren, K.; Stanssens, P.; Baumeister, J.; Holz, J.-B. *Blood* **2011**, *118*, 757–765.
28. The author employed the MeDbz linker for the synthesis of peptides L-**14** and D-**14** to minimize the undesired diacylation at the Dbz linker, to see: Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc.* **2015**, *137*, 7197–7209.
29. Collins, J. M.; Porter, K. A.; Singh, S. K.; Vanier, G. S. *Org. Lett.* **2014**, *16*, 940–943.
30. Subirós-Funosas, R.; El-Faham, A.; Albericio, F. *Tetrahedron* **2011**, *67*, 8595–8606.
31. Annis, I.; Chen, L.; Barany, G. *J. Am. Chem. Soc.* **1998**, *120*, 7226–7238.
32. Kinoshita, S.; Nakakido, M.; Mori, C.; Kuroda, D.; Caaveiro, J. M. M.; Tsumoto, K. *Protein Sci.* **2022**, *31*, e4450.
33. Liu, H.; Schittny, V.; Nash, M. A. *Nano Lett.* **2019**, *19*, 5524–5529.
34. Cortez-Retamozo, V.; Backmann, N.; Senter, P. D.; Wernery, U.; De Baetselier, P.; Muyldermans, S.; Revets, H. *Cancer Res.* **2004**, *64*, 2853–2857.
35. Shah, D. K. *J. Pharmacokinet. Pharmacodyn.* **2015**, *42*, 553–571.
36. Massa, S.; Xavier, C.; Vos, J. D.; Caveliers, V.; Lahoutte, T.; Muyldermans, S.; Devoogdt, N. *Bioconjug. Chem.* **2014**, *25*, 979–988.

37. Chakravarty, R.; Goel, S.; Cai, W. *Theranostics* **2014**, *4*, 386–398.
38. Dehaen, G.; Eliseeva, S. V.; Kimpe, K.; Laurent, S.; Vander Elst, L.; Muller, R. N.; Dahean, W.; Binnemans, K.; Parac-Vogt, T. N. *Chem. - Eur. J.* **2012**, *18*, 293–302.
39. Small differences in the distributions into some organs (for example, higher distribution of L-PMP12A2h1 into bone, but with no significant difference) were observed. Further investigations of the pharmacodynamic properties of L-VHHs and D-VHHs may be needed to understand the underlying mechanism.
40. Ficht, S.; Payne, R. J.; Guy, R. T.; Wong, C.-H. *Chem. - Eur. J.* **2008**, *14*, 3620–3629.
41. Katoh, Y.; Nozaki, S.; Hartanto, D.; Miyano, R.; Nakayama, K. *J. Cell Sci.* **2015**, *128*, 2351–2362.

| | |
|---|---|
| Chapter 1. | Development and Application of a Screening Platform for Mirror-Image Single-Domain Antibodies |
| Section 2. | Identification of a Mirror-Image VHH against Vascular Endothelial Growth Factor |

*Summary: The author developed a novel screening platform to identify an enantiomeric form of the variable domain of the antibody heavy chain (D-VHH) specific for vascular endothelial growth factor A (VEGF-A). The author performed mirror-image screening of two newly constructed synthetic VHH libraries displayed on T7 phage and identified VHH sequences that effectively bound to the mirror-image VEGF-A target (D-VEGF-A). The author subsequently synthesized a D-VHH candidate that preferentially bound the native VEGF-A (L-VEGF-A) with sub-micromolar affinity. Furthermore, immunization studies in mice demonstrated that this D-VHH elicited no anti-drug antibodies (ADAs), unlike its corresponding L-VHH. These findings highlight the utility of this novel D-VHH screening platform in the development of protein therapeutics exhibiting reduced immunogenicity.*

To explore the potential of the enantiomers of the variable domain of the antibody heavy chain (D-VHHs) as therapeutic agents, a mirror-image screening platform has to be developed by first constructing a VHH library. Immune,[1] naïve,[2] and synthetic VHH libraries[3–5] are generally used to identify VHHs with the desired functions. Among these, immune libraries, which are created by directly immunizing camelids with target molecules, are the most commonly used. However, the V-D-J rearrangement, which occurs in the germline VHH sequences from immune and naïve libraries, generates a highly diverse repertoire of complementarity-determining regions (CDR) and framework regions (FRs).[6] Such vast diversity would complicate the synthesis of enantiomeric VHHs from hit sequences derived from the mirror-image screening of immune and naïve libraries. By contrast, using synthetic libraries with common frame regions would be more compatible with mirror-image screening because it would mitigate the potential risk of synthesizing chemically complex hit sequences.[7]

In Chapter 1, Section 1, the author developed a method for generating synthetic VHHs, which the author then used to prepare two model VHHs: the anti-GFP VHH and PMP12A2h1 (the monomeric VHH of caplacizumab).[8] Given that caplacizumab is an approved therapeutic agent with favorable pharmaceutical properties,[9] the author envisioned that PMP12A2h1 would be an appropriate synthetic library frame sequence for the mirror-image screening of VHHs. When preparing VHH libraries, random

elements can be strategically inserted at specific locations in the CDRs of PMP12A2h1, which would then be synthetically accessible. The aim of the present study is therefore to identify functionally interesting D-VHHs by subjecting the newly constructed VHH libraries to mirror-image screening. In this study, the author selected vascular endothelial growth factor A (VEGF-A), an attractive target protein in the treatment of age-related macular degeneration and certain cancers,[10] as a model antigen.

**T7 Phage Synthetic VHH Library Design and Construction**. T7 phage has several advantages over the commonly used filamentous phage, including more rapid cell lysis and less bias toward the amino acids they display.[11] To generate a T7 phage synthetic VHH library, the author employed an efficient library construction technique, which Nonaka *et al.* recently established.[12] Initially, the author constructed a VHH library, in which the random sequence containing NNK codons, whereby N = A/C/G/T and K = G/T, was located only in the CDR3 region. However, no hit sequences were obtained with this set up, probably due to the colossal theoretical diversity (> $10^{30}$) of the designed sequences, which exceeded the technically feasible library size (< $10^{11}$) (data not shown). Therefore, two VHH libraries (VHH library A and VHH library B) were newly designed to strike a balance between theoretical diversity and library size (Figure 1). To limit the diversity of random regions, the author analyzed 526 llama-derived VHH amino acid sequences sourced from the sdAb database.[13] On the basis of the consensus framework sequence of llama-derived VHHs, two amino acids in the frame region were modified (R77N and M78T) so that they differed from those in the caplacizumab scaffold. In VHH library A, the sequences of CDR1 and CDR2 were fixed using the most frequently occurring amino acids identified in the analysis. In VHH library B, two highly variable positions in CDR1 (positions 31 and 33) were randomized, while the other positions in CDR1 and CDR2 were fixed using the same amino acids as those use in VHH library A (Table 1). Given the CDR3 length diversity, the author designed four types of CDR3 sequences with lengths of 9, 12, 15, or 18 amino acids. In VHH library A, CDR3

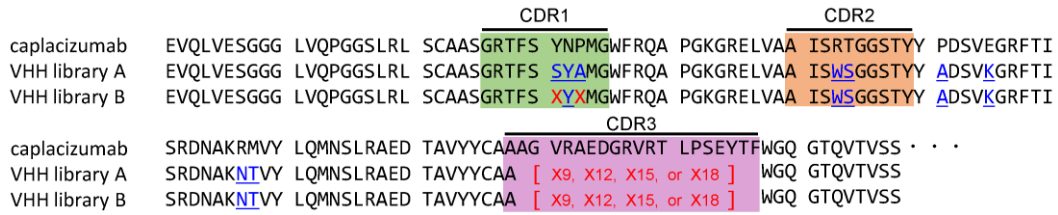


**Figure 1.** Sequences of caplacizumab and the VHH libraries. Parts of CDR1 and CDR3 (red) were randomized with degenerate codons. Underlined residues (blue) indicate amino acids modified using the caplacizumab scaffold as a reference.

randomization was achieved by limiting the diversity of sequences to ~$10^9$ by using degeneration codons (a mixture of nucleotide triplets capable of collectively encoding more than one amino acid[14]) to reproduce the llama-derived sequences (Table 2). In VHH library B, the XYZ codon, which is an improved NNS codon (S = C/G) mimicking the amino acid distribution of natural CDR-H3s,[15] was used for CDR3 randomization.

T7 phages displaying VHH libraries A and B were generated using the published protocol by Nonaka's group.[12] Briefly, library DNA sequences were prepared by ligating the DNA inserts, obtained by overlap extension polymerase chain reaction (OE-PCR) from five double-stranded DNA (dsDNA) fragments, to the 10-3b phage genomic DNA arm (Figure 2A). The resulting products were packaged using a cell-free protein synthesis system.[16] Successful ligation of the OE-PCR products to the T7 phage genomic DNA was confirmed by agarose gel electrophoresis (Figures 2B and 2C). The library sizes were 2.3 × $10^8$ for VHH library A and 4.1 × $10^8$ for VHH library B, which met the sufficient diversity generally required for a successful synthetic library screening.[5] These results showed that the VHH-displaying T7 phage maintained high levels of packaging

**Table 1.** Encoded amino acids at positions 31 and 33 in CDR1 sequence of library B.

| Position | Degenerate Ccdon | Amino acid residues encoded |
|---|---|---|
| 31 | *VNT* | NSTIHRPLDGAV |
| 33 | *HMT/GNT* | NTHPYS/DGAV |
| | | calcd diversity: 1.2 × $10^2$ |

**Table 2.** Encoded amino acids in CDR3 sequence of library A.

| Position | Amino acid resides encoded | | | |
|---|---|---|---|---|
| | X9 | X12 | X15 | X18 |
| 99 | NTSDAGHPRYC | KNRSEDG | NSHRDG | NSDG |
| 100 | TRMAGVSWL | KNRSEDG | KTREAG | RL |
| 101 | NSTIHRPLDGAV | TSPRAG | TRAGPSW | SKA |
| 102 | NSTIHRPLDGAV | TSPRAG | TSPRAG | SRG |
| 103 | TSIPRLAGV | PRLAGV | TSPRAG | NTYS |
| 104 | NTSDAGHPRYC | TIAV | AGVSWL | HPYS |
| 105 | HRPLDGAVYCSF | TRAGSW | NTIYSF | AGV |
| 106 | NSTIHRPLDGAV | TRAGSW | TSIPRL | HPYS |
| 107 | NTIDAVHPLYSF | NSTIYCF | *TSPR* | TSAG |
| 108 | | YSF | TSPRAG | TMSL |
| 109 | | NDY | NSDGYC | SIRL |
| 110 | | NTDAYS | KNRSEDG | HRDG |
| 111 | | | Y | TAPS |
| 112 | | | D | NSDG |
| 113 | | | Y | NSDG |
| 114 | | | | Y |
| 115 | | | | D |
| 116 | | | | Y |
| calcd diversity | 2.2 × $10^9$ | 3.3 × $10^8$ | 3.0 × $10^9$ | 3.0 × $10^8$ |

efficiency, which were similar to those of intact or peptide-displaying T7 phage.[12,16] To confirm the insertion of the VHH sequences, 24 phage clones were randomly picked from each library and the DNA sequences were analyzed. Excluding any frame-shift mutations, 17 clones from VHH library A and 15 clones from VHH library B matched the designed

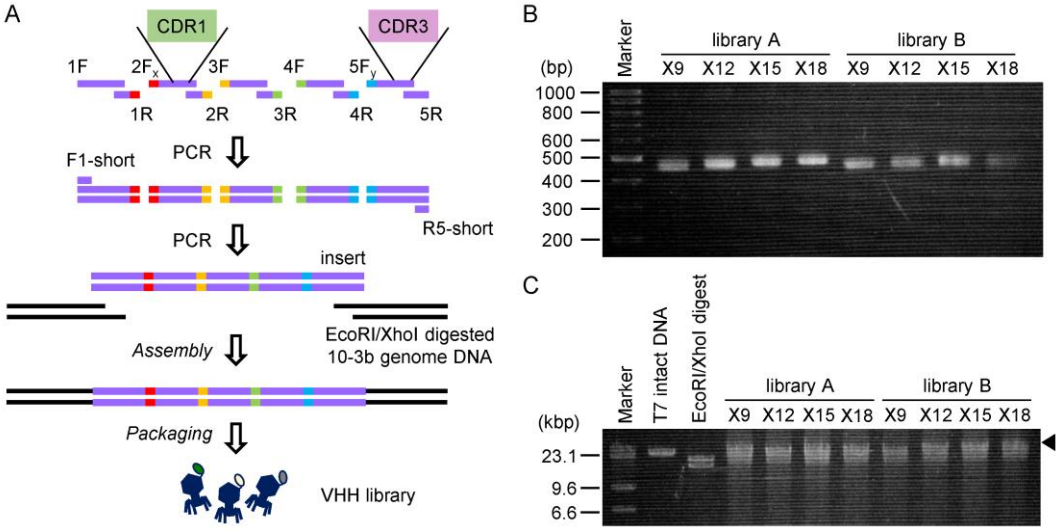


**Figure 2.** Construction of the T7 phage synthetic VHH libraries. (A) Schematic flow chart showing the VHH library construction process. (B) Agarose gel electrophoresis images of overlap-extension PCR products. (C) Agarose gel electrophoresis images of DNA fragments after ligation of VHH inserts and digested T7 phage genomic DNA.

**Table 3.** Representative CDR sequences in the VHH libraries. * indicates the stop codon.

| **Library A** | | | | **Library B** | | | |
|---|---|---|---|---|---|---|---|
| **Clone** | **31** | **33** | **CDR3** | **Clone** | **31** | **33** | **CDR3** |
| A-#1 | S | A | DRGGGLTRTANGYDY | B-#1 | G | N | NGYEACVTPCCNYDY |
| A-#2 | S | A | GRWPALYSRSGRYDY | B-#2 | L | A | PGGNGFESDRW*YDY |
| A-#3 | S | A | HEWAGVYRRGNKYDY | B-#3 | T | P | HDPICHAGEGSLYDY |
| A-#4 | S | A | GRGGYYVSTSSDTGNYDY | B-#4 | A | G | PLGCTETGDAGVYDY |
| A-#5 | S | A | GGRTASFIPAGE | B-#5 | S | G | YYYGVFCHDGERYDY |
| A-#6 | S | A | GTTGRATRTACRYDY | B-#6 | H | V | LRCDDYL*G |
| A-#7 | S | A | TWSDLDYHD | B-#7 | L | P | GHLRGHCYHSYKYDY |
| A-#8 | S | A | NTPSGGITTGYRYDY | B-#8 | S | Y | DGYAGGIGLCGQYDY |
| A-#9 | S | A | DLGGYSGHTSIRSSNYDY | B-#9 | V | N | DCAYDHLEWPTD |
| A-#10 | S | A | NRWSGWITSSGRYDY | B-#10 | L | G | GHKEVAWSVYLRDAVYDY |
| A-#11 | S | A | SDTRVIRTNFYN | B-#11 | P | T | KVRDCRAFRIICGKGYDY |
| A-#12 | S | A | SRSGAAYITADGYDY | B-#12 | I | D | VVYEVGHSA |
| A-#13 | S | A | GRRGYSASASLHSSGYDY | B-#13 | G | P | GAHARHGGY |
| A-#14 | S | A | DERRSLIRTSCKYDY | B-#14 | S | V | HGYPWGSSV |
| A-#15 | S | A | NERGRWFLRGGGYDY | B-#15 | G | S | DGGGCRYDY |
| A-#16 | S | A | DESGGVFTRSNDYDY | | | | |
| A-#17 | S | A | DAAGGVSSSGYGYDY | | | | |

VHH sequences (Table 3). The author also confirmed that both libraries contained VHHs with a variety of CDR3 lengths (i.e., 9, 12, 15, or 18 amino acids). Next, the sequences of the prepared libraries were characterized using next-generation sequencing (NGS). The diversity index (Pielou's *J*' value)[17] in each library was > 0.994, indicating that the CDR3 amino acid sequences were evenly distributed (Figure 3).

To validate the VHH libraries A and B, the author screened them against green fluorescent protein (GFP), lymphocyte activation gene-3 (LAG-3), and VEGF-A as model targets. After five rounds of biopanning plates presenting the immobilized targets, several phage clones were randomly selected and their DNA and corresponding peptide sequences were analyzed. For GFP and LAG-3, enrichment of sequences with high homology was observed in VHH library A (Table 4). For VEGF-A, a specific sequence E1 (CDR3 sequence [VDVYYVYGGE]) was enriched from T7 VHH library B. These

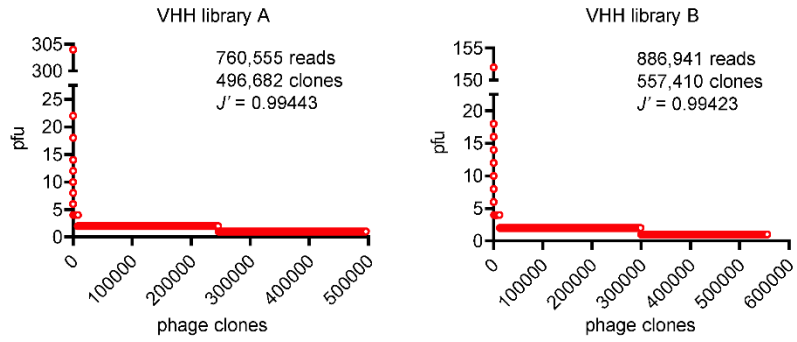


**Figure 3.** Comprehensive assessment of CDR3 sequences in VHH library A and VHH library B by next-generation sequencing (NGS). Total reads and the number of clones in each library are shown. Pielou (*J*') indexes were calculated as described before.[17]

**Table 4.** VHH sequences of the selected L-target-binding phage. Blue letters: common sequence. VHH[E1] was recombinantly expressed for further analysis.

| VHH | Library | Target | 31 | 33 | CDR3 | Ratio |
|---|---|---|---|---|---|---|
| C1 | A | L-GFP | S | A | GKRGPTGGYFYT | 3/6 |
| C2 | A | L-GFP | S | A | GKRGGTGRYFYT | 1/6 |
| C3 | A | L-GFP | S | A | GERALTGAYFYT | 1/6 |
| C4 | A | L-GFP | S | A | RLASRFY | 1/6 |
| D1 | A | L-LAG-3 | S | A | SDGRGVWWYFDD | 2/6 |
| D2 | A | L-LAG-3 | S | A | SNGSGTWWYYDD | 2/6 |
| D3 | A | L-LAG-3 | S | A | SDGRGAWWYYNA | 1/6 |
| D4 | A | L-LAG-3 | S | A | GESRTVIIRTGGYDY | 1/6 |
| E1 | B | L-VEGF | S | D | VDVYVYGGE | 6/12 |
| E2 | B | L-VEGF | V | N | RYHHGGGIL | 2/12 |
| E3 | B | L-VEGF | H | V | FHTSYRS | 1/12 |
| E4 | B | L-VEGF | N | T | WCVMSGDTFVGVYDY | 1/12 |
| E5 | B | L-VEGF | L | D | VHHHGHVFT | 1/12 |
| E6 | B | L-VEGF | V | V | SRFHPRGLS | 1/12 |

results suggest that the author had successfully constructed a system that could be used to identify VHHs against various targets.

**Mirror-Image Phage Display Selection.** To explore D-VHHs with target binding capacity, the author selected VEGF-A as a model target for mirror-image screening (Figure 4). The mirror-image form of the active VEGF-A protein (D-VEGF-A) was synthesized as previously described.[18] This time, however, an additional poly-histidine (His)-tag was attached to the D-VEGF-A N-terminus via two glycine residues, which enabled its adhesion to the nickel (Ni) plate (Figure 5). Peptide fragments **1**, **2**, with an *N*-acyl-*N'*-methyl-benzimidazolinone (MeNbz) linker at the C-terminus,[19] and C-terminal peptide **3** were generated using Fmoc-solid phase peptide synthesis (Fmoc-SPPS). Three unprotected peptide fragments were condensed by one-pot native chemical ligation (NCL).[20] Subsequently, the resulting full-length protein **4** was oxidized to form intra- and intermolecular disulfide bonds,[18] yielding the folded D-VEGF-A (protein **5**).

With a functional D-VEGF-A in hand, the author performed mirror-image screening using the T7 VHH libraries. After four rounds of biopanning the D-VEGF-A-immobilized plate with VHH libraries A and B, the sequences of the enriched clones were ranked using NGS analysis (Table 5). The highly homologous VHH[A2] and VHH[A4] were identified through VHH library A screening. In addition, library B screening revealed marked VHH[B1] sequence enrichment. Among the top ranked VHH sequences, some (e.g., VHH[A1], VHH[B2], and VHH[B3]) contained several histidine residues in their CDR3.

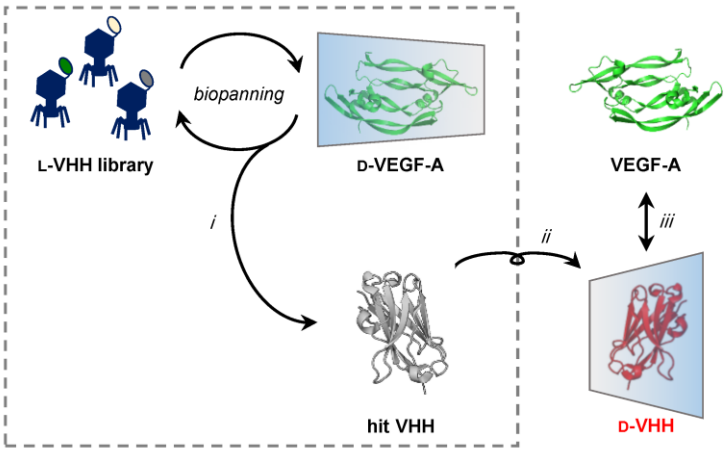


**Figure 4.** Scheme for selection of VEGF-A-binding D-VHHs using mirror-image T7 phage display. (i) The VHH library is used in the biopanning of chemically synthesized D-VEGF-A. (ii) The identified VHH sequence is then chemically synthesized from D-amino acids. (iii) The ability of the candidate D-VHH to bind L-VEGF-A is assessed. Protein data bank (PDB) ID: 3QTK [VEGF-A]; 5TP3 [VHH].

The VHH sequences containing multiple histidine and arginine residues within their CDR3s were excluded from further analysis as they could form complexes with $Ni^{2+}$.[21,22] Phages displaying such sequences could directly bind to the Ni plate but not to D-VEGF-A. Monitoring the number of phages displaying VHH[A2] and VHH[B1] confirmed that both VHH types were enriched in each round (Figure 6A). To further evaluate the binding selectivity of the candidate sequences to D-VEGF-A, individual phage binding assays were performed against D-VEGF-A or L-VEGF-A. Phage displaying VHH[A2], VHH[A3], and VHH[B1] had a > 400-fold higher preference for D-VEGF-A than L-VEGF-A (Figure 6B). The author also performed binding assays with VHH[E1]-displaying phage identified in the screening against L-VEGF-A and found that the

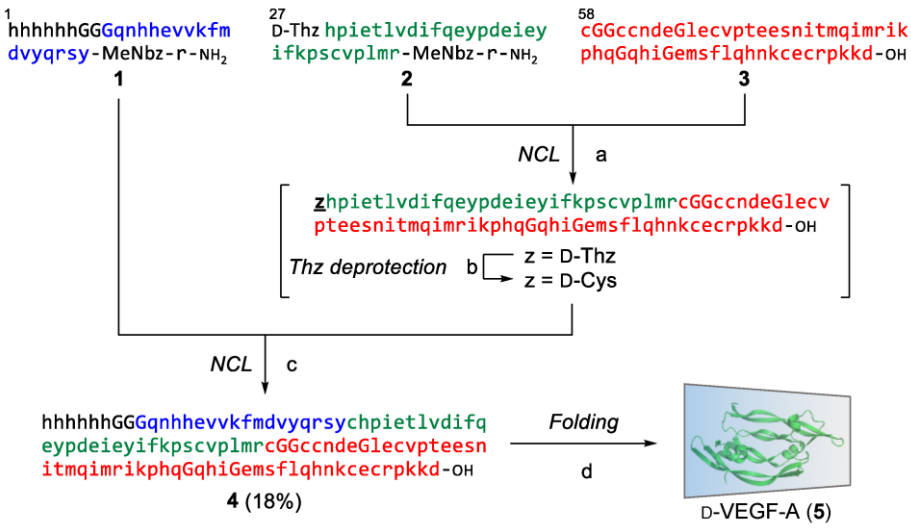


**Figure 5.** Synthesis of D-VEGF-A. *Reagents and conditions*: (a) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (b) methoxyamine; (c) NaOH (adjustment pH to 7) (d) 6 M guanidine·HCl, 100 mM Tris (pH 8.4), then 40-fold dilution into 2.0 mM reduced glutathione, 0.40 mM oxidized glutathione, and 100 mM Tris (pH 8.4).

**Table 5.** VHH sequences of the selected D-VEGF-binding phage. Blue letters: common sequence. Red letters: histidine or arginine. VHH[A2], VHH[A3], and VHH[B1] were recombinantly expressed for further analysis.

| VHH | Library | Target | 31 | 33 | CDR3 | NGS count |
|---|---|---|---|---|---|---|
| A1 | A | D-VEGF | S | A | YGHRRRVRY | 111933 |
| A2 | A | D-VEGF | S | A | EGPGLAWWSYDA | 33460 |
| A3 | A | D-VEGF | S | A | RTGGAGFTRRYGYDY | 26667 |
| A4 | A | D-VEGF | S | A | EGPGVAWWNYDT | 26444 |
| B1 | B | D-VEGF | V | H | IADRCYDTC | 133374 |
| B2 | B | D-VEGF | G | V | YHHEHVEDTMFWYDY | 59356 |
| B3 | B | D-VEGF | I | H | VSRSHPHFY | 16494 |

VHH[E1] clone bound L-VEGF-A with a high degree of selectivity.

To determine the binding affinity of the candidate VHHs, four recombinant VHHs (i.e., VHH[A2], VHH[A3], VHH[B1], and VHH[E1]) were prepared using the *Escherichia coli* expression system. Surface plasmon resonance (SPR) analysis showed that VHH[A2], VHH[A3], and VHH[B1] bound D-VEGF-A with moderate affinity (VHH[A2]: $K_D$ = 396 ± 13 nM; VHH[A3]: $K_D$ = 351 ± 4 nM; VHH[B1]: $K_D$ = 66 ± 9 nM); however, they did not bind to L-VEGF-A (Figure 7). By contrast, VHH[E1] bound to L-VEGF-A with an affinity of 185 ± 5 nM but not to D-VEGF-A.

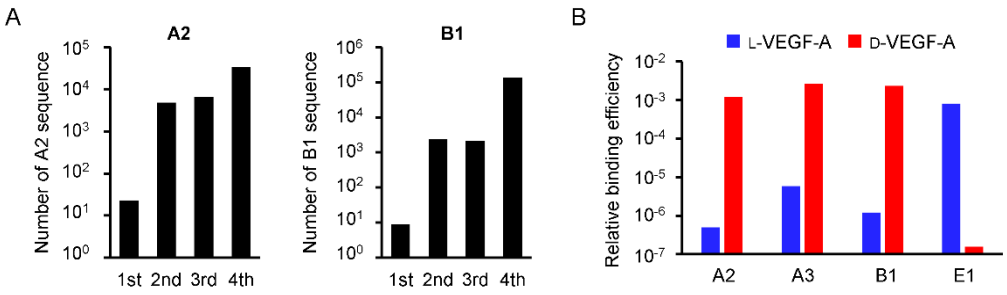


**Figure 6.** Enrichment and binding of selected D-VEGF-binding phage. (A) Enrichment of A2 phage or B1 phage after four rounds of biopanning. The histogram shows the number of A2 phage or B1 phage identified by NGS. (B) Phage binding assays with L-/D-VEGF-A.

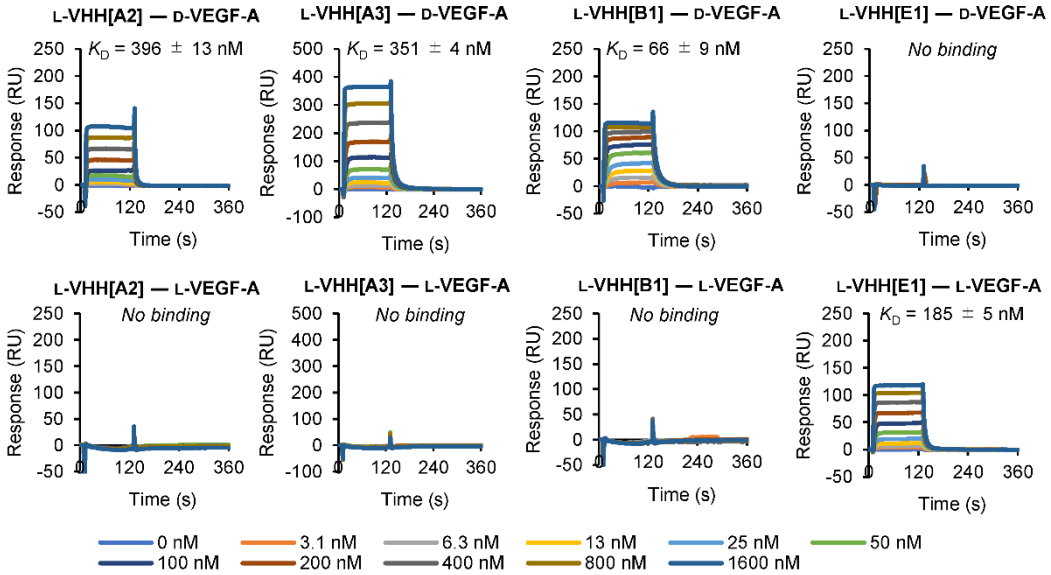


**Figure 7.** Representative SPR curves showing the binding of recombinant VHHs to enantiomeric forms of VEGF-A. $K_D$ values (mean ± SD) were determined from the saturation curves generated from triplicate experiments.

**Synthesis and Biological Evaluation of D-VHH.** The author selected VHH[B1], which bound to D-VEGF-A with the highest affinity, as a hit sequence for further investigation to obtain a L-VEGF-A-binding D-VHH. The mirror-image protein with the same sequence, albeit composed of D-amino acids (D-VHH[B1]), was chemically synthesized as described in Chapter 1, Section 1[8] but with some modifications (Figure 8).

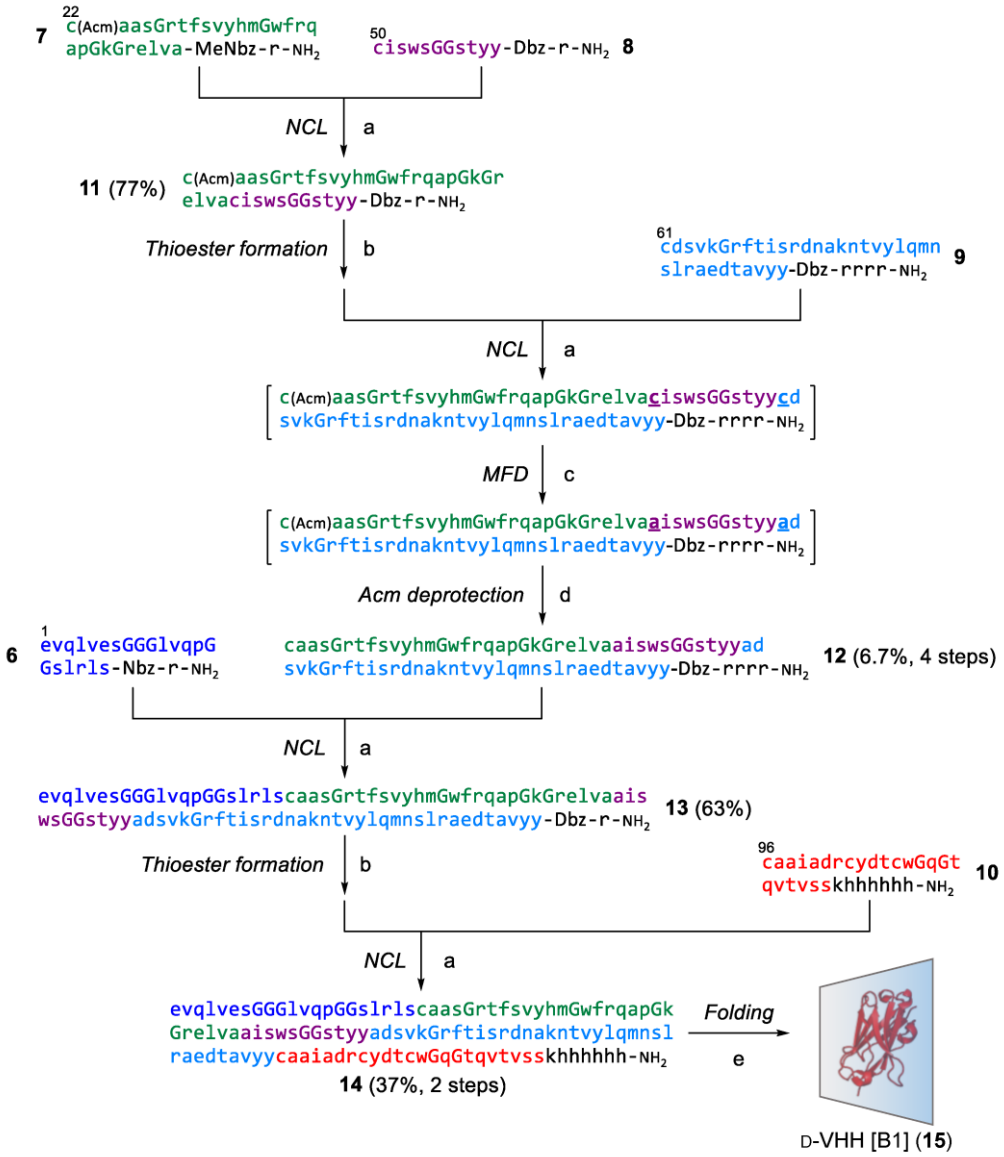


**Figure 8.** Synthesis of D-VHH[B1]. *Reagents and conditions*: (a) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (b) $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, then MPAA and TCEP; (c) VA-044, TCEP, MESNa, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 6.5); (d) $PdCl_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, and then DTT; (e) 6 M guanidine·HCl, 20 mM phosphate (pH 7.5), then gradually dialyzed against PBS (pH 7.4).

A D-$His_6$-tag was attached to the C-terminus of the VHH sequence to improve solubility during the purification by high-performance liquid chromatography (HPLC). Because a full-length, 125-amino-acid sequences is difficult to construct by conventional Fmoc-SPPS, the author used sequential NCLs with two cysteine residues ($Cys^{22}$, $Cys^{96}$) in the frame regions. To prepare the 74-residue middle fragment, $Ala^{50}$ and $Ala^{61}$ were employed as additional NCL sites, considering the coupling efficiency of the peptides during the Fmoc-SPPS stage and the solubility of the resulting product during the purification stage.

The five peptide fragments with appropriate linkers and auxiliary arginines at the C-terminus (**6**–**10**) were synthesized by Fmoc-SPPS. Peptide **7** and peptide **8** were ligated by a 4-mercaptophenylacetic acid (MPAA)-mediated NCL reaction to obtain peptide **11**. $NaNO_2$-mediated activation of the C-terminal Dbz linker in **11**,[23] followed by NCL with peptide **9** in the presence of MPAA, yielded the middle fragment precursor. Subsequently, desulfurization[24] and deprotection of the acetamidomethyl (Acm) group[25] were performed to give rise to middle fragment **12**. Next, **12** was ligated with **6** in the presence of MPAA to yield peptide **13**. The final ligation between **13** and the C-terminal fragment **10** gave rise to the desired D-VHH[B1] **14**. Full-length D-VHH **14** was folded by dialysis in the presence of glutathione. Protein aggregation was observed during folding; however, using a lower protein concentration (~0.25 mg/mL) improved the yield of folded protein **15** to approximately 37%.

The secondary structure of the folded D-VHH[B1] was assessed by circular dichroism (CD), which showed maximum absorption at ~215 nm and minimum absorption at ~200 nm (Figure 9A). The spectrum of D-VHH[B1] mirrored that of the

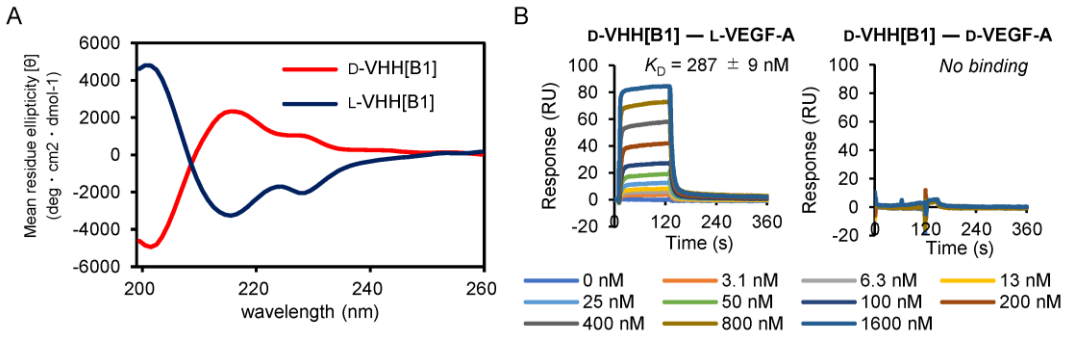


**Figure 9.** Structure and binding activity of D-VHH[B1]. (A) CD spectra of recombinant L-VHH[B1] and synthetic D-VHH[B1]. (B) Representative SPR curves showing the interaction between the chemically synthesized D-VHH[B1] and the enantiomeric forms of VEGF-A. $K_D$ values (mean ± SD) were determined from the saturation curves generated from triplicate experiments.

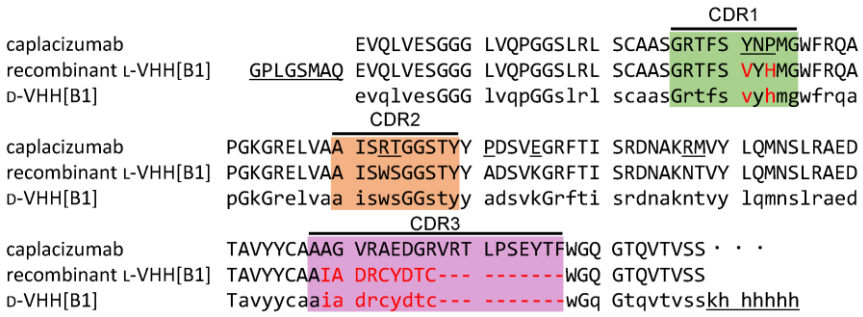


**Figure 10.** Sequences of the caplacizumab, recombinant L-VHH[B1], and synthetic D-VHH[B1]. Capital letters denote L-amino acids. Lowercase letters denote D-amino acids.

recombinant L-VHH[B1] prepared by expression in *E. coli*. SPR was used to evaluate the binding affinity of synthetic D-VHH[B1] for L-VEGF-A and D-VEGF-A. D-VHH[B1] (at 3.1–1,600 nM) was applied to biotinylated L-VEGF-A or D-VEGF-A, which was immobilized on streptavidin-coated or CM5 sensor chips, respectively. D-VHH[B1] bound to L-VEGF-A ($K_D$ = 287 ± 9 nM) but not to D-VEGF-A (Figure 9B). These results showed that the binding activity was reproduced with the mirror-image combination; however, the affinity of the synthetic D-VHH[B1] for L-VEGF-A was slightly lower than that of the recombinant L-VHH[B1] for D-VEGF-A (Figures 7 and 9B). The lower affinity of D-VHH[B1] could be attributed to the difference in N- and C-terminal sequences between the synthetic and recombinant forms of VHH (Figure 10).

**Immunogenicity of VEGF-A-binding D-VHH.** One advantage of D-VHHs over L-VHHs is that they are less likely to induce unwanted anti-drug antibody (ADA) production.[8] Thus, the author next assessed the immunogenicity of the VEGF-A-binding D-VHH[B1] in mice. The recombinantly expressed VEGF-A-binding L-VHH[E1] was selected as a control L-VHH, and its purity was confirmed to be similar to that of synthetic D-VHH[B1] (Figure 11A). D-VHH[B1] and L-VHH[E1] (25 μg of each) were suspended in a strong adjuvant and administered intraperitoneally to the BALB/c mice. To assess the immune response during long-term VHH administration, boosting immunizations were performed on days 14 and 28. The sera of immunized mice were collected on days 14, 28, and 42 after the first injection and used to evaluate the generation of ADAs by enzyme-linked immunosorbent assay (ELISA). In the L-VHH[E1]-treated group, low ADA levels were observed after the first immunization; however, this level increased after boosting (Figure 11B). On day 42, four out of five L-VHH-treated mice had high L-VHH-binding antibody titers, which were dependent on the concentration of the coating antigen (Figure 11C). By contrast, even on day 42, no ADAs against D-VHH[B1] were

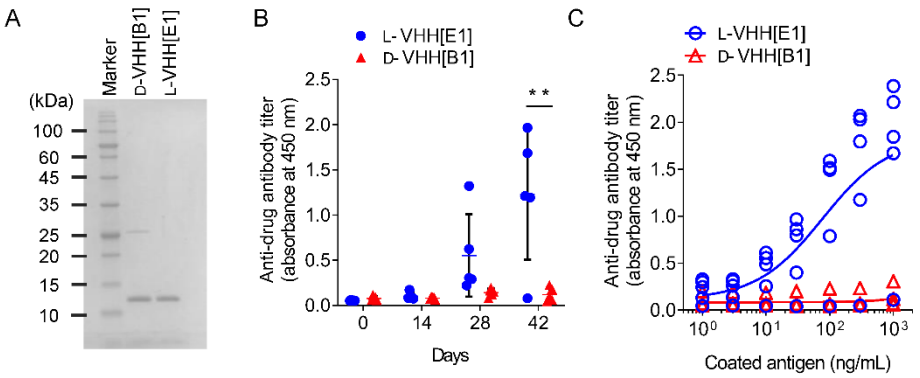


**Figure 11.** Immunogenicity of VEGF-A-binding L-VHH and D-VHH. (A) CBB staining of D-VHH[B1] and L-VHH[E1]. (B) Anti-VHH antibody titers were measured in mouse sera on days 0, 14, 28, and 42 after the intraperitoneal injection of L-VHH[E1] or D-VHH[B1] (25 µg/injection). $N = 5$ mice per group. Data are presented as the mean ± SD. (ns: not significant; **** $p < 0.0001$; Two-way ANOVA followed by Sidak's multiple comparison test.) (C) Detection of ADA in mouse sera (day 42) from intraperitoneal injection of L-VHH[E1] or D-VHH[B1] by ELISA with varying concentrations of the plate-coating antigen.

detected in the sera of the D-VHH[B1]-treated mice; these results align with the conclusions of the previous immunogenicity assessment of two model VHH enantiomers described in Chapter 1, Section 1.[8]

In summary, the author demonstrated that two T7 phage synthetic VHH libraries could be successfully used in mirror-image screening to identify novel D-VHH therapeutic candidates. The VHH libraries were compatible with chemical synthesis, during which random residues containing degeneration codons were planted within the CDR1 and CDR3 sequences using a modified caplacizumab framework. Biopanning of the D-VEGF-A target led to the identification of a VEGF-A-binding D-VHH; its hit sequence was then chemically synthesized by stepwise NCLs. Importantly, although the resulting D-VHH was selective for its therapeutic target, it did not induce ADAs in mice. Thus, the author envisage that the D-VHH discovery platform presented herein will serve as a foundation for the future application of D-VHHs as therapeutic agents and diagnostic tools.

Experimental Section

**Bacterial and Bacteriophage Strains and Culture.** The *Escherichia coli* strain BLT5403 [$F^-$, *ompT*, *hsdS*$_B$ ($r_B^- m_B^-$), *gal*, *dcm* pAR5403 (*Amp*$^R$)] and the bacteriophage T7Select 10-3b were obtained from Merck Millipore (Darmstadt, Germany). The pAR5403 plasmid was isolated from BLT5403 cells using the FastGene Plasmid Mini Kit (NIPPON Genetics Co., Ltd., Japan) according to the manufacturer's instructions. The *E. coli* strain SHuffle (New England BioLabs, Ipswich, MA, USA) harboring pAR5403 (SHuffle5403) was used to amplify the VHH phages. The *E. coli* strains BLT5403 and SHuffle5403 were cultured in LB medium (Nacalai Tesque, Inc., Kyoto, Japan) containing 100 µg/mL ampicillin sodium salt (Nacalai Tesque, Inc.).

**Phage Titration.** Phage libraries or phage pools were diluted in LB medium or SM Buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM $MgSO_4$, 0.01% gelatin) as required. The BLT5403 cells (in exponential growth phase) and the diluted phages were suspended in 3–4 mL of LB top agar (0.75%) and plated onto Petri dishes. The plates were incubated at 37 °C and the number of plaques was counted.

**Overlap-Extension PCR.** All oligonucleotides used to construct VHH libraries are listed in Table 6. The VHH DNA libraries were constructed by overlap-extension PCR (OE-PCR) using PrimeStar HS DNA Polymerase (Takara Bio Inc., Shiga, Japan), as illustrated in Figure 2A. The first round involved the extension of five fragments (1F-1R, $2F_x$-2R [x = a, b1, or b2], 3F-3R, 4F-4R, $5F_y$-5R [y = a1, a2, a3, a4, b1, b2, b3, or b4]) with two corresponding oligonucleotides (1 µM of each) as primers. The first round of PCR was conducted as follows: a 10 sec initial denaturation at 96 °C; five cycles of 96 °C for 10 sec, 55 °C for 10 sec, and 72 °C for 10 sec; and a final step of 10 sec at 72 °C, followed by cooling to 4 °C. After the first round of PCR, the five PCR products (10 µL of each) were mixed and annealed as follows: a 10 sec initial denaturation at 96 °C; 20 cycles of 96 °C for 10 sec, 60 °C for 10 sec, and 72 °C for 10 sec; and a final step of 10 sec at 72 °C. The second round of PCR was performed using the F1-short and R5-short oligonucleotides as primers and each PCR mixture as the template. The second round of PCRs was conducted as follows: a 30 sec initial denaturation at 98 °C; 10 cycles of 98 °C for 10 sec and 68 °C for 60 sec; followed by cooling to 4 °C. The products were purified using the FastGene Gel/PCR Extraction Kit (NIPPON Genetics Co., Ltd., Japan). The final OE-PCR products were analyzed by electrophoresis on a 2.4% agarose gel.

**Table 6.** List of oligonucleotides used for library construction.

| Name | DNA (5'-3') |
|---|---|
| 1F | GTCAGGTGTGATGCTCGGGGATCCGAATTCTGGTGGAGGCGGTTCAGGCG GAGGTGGCTCCGGCGGTGGAGGATCTGAAGTTCAGCTGGTT |
| 1R | CAGACGCAGAGAACCACCAGGTTGTACCAGACCACCACCAGATTCAACCA GCTGAACTTC |
| $2F_a$ | GGTTCTCTGCGTCTGTCTTGTGCAGCATCTGGTCGCACCTTCAGCAGCTAT GCGATGGGCTGGTTTCGTCAGGCTCCAGG |
| $2F_{b1}$ | GGTTCTCTGCGTCTGTCTTGTGCAGCATCTGGTCGCACCTTCAGC**VNT**TAT **HMT**ATGGGCTGGTTTCGTCAGGCTCCAGG |
| $2F_{b2}$ | GGTTCTCTGCGTCTGTCTTGTGCAGCATCTGGTCGCACCTTCAGC**VNT**TAT **GNT**ATGGGCTGGTTTCGTCAGGCTCCAGG |
| 2R | AGTTCACGACCTTTGCCTGGAGCCTGACGA |
| 3F | CAAAGGTCGTGAACTGGTGGCCGCGATCAGCTGGAGCGGCGGCAGCACCT ATTACGCCGATTCCGTAAAAGGTCGTTTC |
| 3R | GTCCCGGGAGATGGTGAAACGACCTTTTAC |
| 4F | ACCATCTCCCGGGACAACGCGAAGAACACCGTATACCTGCAGATGAACAG CCTGCGTGC |
| 4R | ACGGCCGTGTCCTCTGCACGCAGGCTGTTC |
| $5F_{a1}$ | AGAGGACACGGCCGTCTATTACTGTGCTGCT**NVTDBGVNTVNCVBTNVCB NTVNCNHT**TGGGGCCAGGGTACGCAGGTCAC |
| $5F_{a2}$ | AGAGGACACGGCCGTCTATTACTGTGCTGCT**RRWRRWVSTVSCSBGRYTD SGDSGWMTTHCDATDMC**TGGGGCCAGGGTACGCAGGTCAC |
| $5F_{a3}$ | AGAGGACACGGCCGTCTATTACTGTGCTGCT**VRTRVGNSGVSTVSCKBGW HTMBCMSTVSCDRTRRM**TATGATTATTGGGGCCAGGGTACGCAGGTCAC |
| $5F_{a4}$ | AGAGGACACGGCCGTCTATTACTGTGCTGCT**RRTCKCSKAVGTWMCYMT GBCYMTRSCWYGMKTSRCNCTRRCRRT**TATGATTATTGGGGCCAGGGTA CGCAGGTCAC |
| $5F_{b1}$ | AGAGGACACGGCCGTCTATTACTGTGCTGCT***XYZXYZXYZXYZXYZXYZXYZXYZ XYZXYZ***TGGGGCCAGGGTACGCAGGTCAC |
| $5F_{b2}$ | AGAGGACACGGCCGTCTATTACTGTGCTGCT***XYZXYZXYZXYZXYZXYZXYZXYZ XYZXYZXYZXYZXYZ***TGGGGCCAGGGTACGCAGGTCAC |
| $5F_{b3}$ | AGAGGACACGGCCGTCTATTACTGTGCTGCT***XYZXYZXYZXYZXYZXYZXYZXYZ XYZXYZXYZXYZXYZ***TATGATTATTGGGGCCAGGGTACGCAGGTCAC |
| $5F_{b4}$ | AGAGGACACGGCCGTCTATTACTGTGCTGCT***XYZXYZXYZXYZXYZXYZXYZXYZ XYZXYZXYZXYZXYZXYZXYZXYZ***TATGATTATTGGGGCCAGGGTACGCAGG TCAC |
| 5R | AGGCCCCAAGGGGTTAACTAGTTACTCGAGTTAAGAGCTAACGGTGACCT GCGTACCC |
| F1-short | GTCAGGTGTGATGCTCGGGGATCCGAATTC |
| R5-short | AGGCCCCAAGGGGTTAACTAGTTACTCGAG |

DNA degeneracy is represented by the IUB code (R = A/G, Y = C/T, M = A/C, K = G/T, S = C/G, W = A/T, H = A/C/G, B = C/G/T, V = A/C/G, D = A/G/T, N = A/C/G/T). XYZ codon[15] (italic) contains unequal nucleotide ratios at each position of the codon triplet. X contains 19% A, 17% C, 38 % G, and 26 % T; Y contains 34 % A, 18% C, 31% G, and 17% T; and Z contains 76% C and 24% G. Degenerate codons are shown in bold text.

**Construction of VHH Libraries.** The T7Select 10-3b vector (Merck Millipore) was digested with EcoRI and XhoI (both from New England Biolabs). The OE-PCR fragments and the EcoRI/XhoI digested T7Select 10-3b vector arms were assembled using the NEBuilder HiFi DNA Assembly Master Mix (New England BioLabs) and purified using MonoFas isolation systems (GL Science, Japan) according to the manufacturer's specifications. The assembled DNA constructs were analyzed by electrophoresis on a 1.0% agarose gel. The mixture of assembled phage DNA constructs (x9:x12:x15:x18 = 20:26:66:41, where 'x' refers to the number of amino acids in the CDR3; final 1 nM), pAR5403 (final 1 nM), and PEG 8000 (final 2%) were added to 9 µL of LS70 Master Mix, a component of the myTXTL T7 Expression kit (Daicel Arbor Biosciences, Ann Arbor, MI), to a final reaction volume of 12 µL; this reaction was then incubated at 29 °C for 16 h. The titers of the prepared phage libraries were assessed as described above. The phages were then amplified in 500 mL of liquid medium containing SHuffle5403 until bacterial lysis occurred. The amplified libraries were centrifuged at 4 °C, 8,000 × *g*, for 30 min and the supernatants were filtered by Nalgene™ Rapid-Flow™ Sterile Disposable Filter Units with PES (0.45 µm, 500 mL; Thermo Fisher Scientific, Waltham, MA, USA). The libraries were stored as 10% glycerol stocks at –80 °C. Quality control was performed by sequencing 24 clones from each library. The diversity of each library was analyzed using the Illumina iSeq 100 sequencing system (Illumina, San Diego, CA, USA) as previously described.[12]

**Phage Display Selection.** His-tagged GFP (cat# 13105-S07E) and LAG3 (cat# 16498-H08H) were purchased from SinoBiological (Japan). His-tagged human VEGF (121 aa) was purchased from ProSpec-Tany TechnoGene Ltd. (Israel). Before each selection round, the VHH libraries were concentrated as follows. 2.5 mL of 5 M NaCl solution was added to 10 mL of the amplified phage library, and the mixture was incubated on ice for 15 min. After centrifugation at 4 °C, 12,000 × *g*, for 15 min, the supernatants were mixed with 12.5 mL of 20% PEG8000 and incubated on ice for 1 h. The PEG-precipitated solutions were centrifuged at 4 °C, 12,000 × *g*, for 15 min and the supernatants were discarded. The pellets were resuspended with 100 µL of PBS (before round 1), or 1 mL of PBS (before rounds 2, 3, 4, and 5). In rounds 1 and 3, Ni-coated plates (Thermo Fischer Scientific) were treated for 30 min with His-tagged target proteins in TBS (100 µL/well, round 1: 1.0 µg/well; round 3: 0.25 µg/well). Next, the wells were washed three times with PBS and blocked with 4% Blockace (MEGMILK SNOW BRAND, 200 µL/well) for 2 h. After three washes with PBS containing 0.05% Tween 20 (PBST), the VHH libraries were suspended in PBS containing with 1% Blockace and 0.05% Tween 20 (100 µL/well), applied to the coated wells, and incubated for 10 min. After 10 washes with

PBST, the bound phages were eluted from the wells with 500 mM imidazole in TBS (100 µL/well) for 1 min. The eluted phages were amplified by incubating with SHuffle5403 *E. coli* in 10 mL of LB medium until bacterial lysis occurred. The amplified phages were then used in the next round of biopanning. In rounds 2 and 4, MaxiSorp plates (Thermo Fisher Scientific) were coated with His-tagged target proteins in 50 mM sodium carbonate buffer (pH 9.4; 100 µL/well, round 2: 0.5 µg/well; round 4: 0.25 µg/well) at 4 °C overnight. After coating, the wells were washed three times with PBS and blocked with PBS containing 3% BSA (200 µL/well) for 2 h. After three washes with PBST, the VHH libraries were suspended in PBS containing 3% bovine serum albumin (BSA) and 0.05% Tween 20 (100 µL/well), applied to the coated wells, and incubated for 10 min. After 10 washes with PBST, the bound phages were eluted by incubating them with fresh SHuffle5403 *E. coli* (100 µL/well) for 2 min. An additional elution step involved replacing the liquid in the wells with 5 M NaCl solution and incubating for 1 min. The eluted phages were amplified by incubating with SHuffle5403 *E. coli* in 10 mL LB medium until bacterial lysis occurred. When screening against GFP, VEGF-A, and LAG3, six individual clones (from round 5) were randomly selected and sequenced. When screening against D-VEGF-A, the target-binding sequences were analyzed using the Illumina iSeq 100 sequencing system (Illumina).

**NGS Data Analysis.** Fastq files generated by the iSeq 100 sequencing system were analyzed using Python 3.9.12 scripts. Initially, the sequence reads were translated to amino acid sequences. The sequences proceeding the g10 capsid protein sequence were extracted as VHH sequences. To reconstruct the complete VHH sequence, the forward read (containing the CDR1 and CDR2 sequences) and the reverse read (containing the CDR3 sequence) were concatenated using the coordinate information assigned to each sequence read. The read counts of the CDR3 sequences were then calculated and arranged in descending order of frequency. Shannon's diversity index was calculated to evaluate the diversity within each library as previously described.[12] Next, the diversity index *J'* (as known as Pielou's evenness index) was calculated by dividing the logarithm of Shannon's diversity indexes by the logarithm of the number of samples;[17] an increase in *J'* from 0 to 1 indicated increasing diversity. Graphs were drawn using GraphPad Prism software.

**Phage Binding Assay.** Ni-coated plates (Thermo Fisher Scientific) were treated for 30 min with His-tagged D-VEGF-A in TBS (100 µL/well, 0.25 µg/well). Next, the wells were washed three times with PBS and blocked with PBS containing 3% BSA (200 µL/well) for 2 h. After three washes with PBST, individual phages were suspended in PBS containing 3% BSA and 0.05% Tween 20 (100 µL/well), applied to the wells and

incubated for 30 min. After 10 washes with PBST, the bound phages were eluted by incubating with 500 mM imidazole in TBS (100 µL/well) for 1 min. Finally, the titers of the eluted phages were measured.

**General Procedure of Peptide Synthesis.** All reagents and solvents were purchased from Watanabe Chemical Industries, Ltd. (Hiroshima, Japan), Kokusan Chemical Industries, Ltd. (Kanagawa, Japan), Sigma-Aldrich JAPAN (Tokyo, Japan), Wako Pure Chemical Industries, Ltd. (Osaka, Japan), Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan) or Nacalai Tesque, Inc. (Kyoto, Japan). For analytical high-performance liquid chromatography (HPLC), a Cosmosil 5C18-AR300 column (4.6 × 250 mm, Nacalai Tesque, Inc.) or a Cosmosil 5C4-AR300 column (4.6 × 150 mm, Nacalai Tesque, Inc.) were employed with a linear gradient of $CH_3CN$ containing 0.05% (v/v) trifluoroacetic acid (TFA) at a flow rate of 1 mL/min (25 °C or 40 °C). The products were detected by UV absorbance at 220 nm. For preparative HPLC, a Cosmosil 5C18-AR300 column (20 × 250 mm, Nacalai Tesque, Inc.) or a Cosmosil 5C4-AR300 column (20 × 150 mm, Nacalai Tesque, Inc.) was employed with a linear gradient of $CH_3CN$ containing 0.05% TFA at a flow rate of 8 mL/min (room temperature). All peptides were characterized by ESI-MS (LCMS-2020, Shimadzu or Micromass ZQ, Waters). Standard Fmoc-based solid-phase peptide synthesis (Fmoc-SPPS) was performed using an automatic peptide synthesizer (PSSM-8, Shimadzu). The following side chain protected amino acids were employed: Arg(Pbf), Asn(Trt), Asp(O*t*-Bu), Cys(Trt), Gln(Trt), Glu(O*t*-Bu), His(Trt), Lys(Boc), Ser(*t*-Bu), Thr(*t*-Bu), Trp(Boc), and Tyr(*t*-Bu). Fmoc protected amino acids (5 equiv) were coupled using *N*,*N'*-diisopropylcarbodiimide (DIC) (5 equiv), and ethyl cyano(hydroxyimino)acetate (Oxyma Pure) (5 equiv) in *N*,*N*-dimethylformamide (DMF) for 60 min twice. Fmoc protection group removal was performed by incubating with 20% piperidine in DMF for 4 min twice. Microwave-assisted peptide synthesis by Fmoc-SPPS was performed on a Liberty BLUE (CEM Japan) peptide synthesizer. The peptide chain was elongated using standard protocols.[8,26]

**D-VEGF-A$^{1–26}$ (1).** C-terminal arginine tag was manually loaded on NovaSyn TGR resin (400 mg, 0.10 mmol) by Fmoc-D-Arg(Pbf)-OH (320 mg, 0.50 mmol), ethyl cyano(hydroxyimino)acetate (OxymaPure) (71 mg, 0.50 mmol) and *N*,*N'*-diisopropylcarbodiimide (DIC) (160 µL, 1.0 mmol) in *N*,*N*-dimethylformamide (DMF) (1.8 mL) for 2 h. *N*-acyl-*N'*-methyl-benzimidazolinone (MeNbz) linker was then loaded on the resin by Fmoc-MeDbz-OH (190 mg, 0.50 mmol), 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) (190 mg, 0.50 mmol), 1-hydroxybenzotriazole monohydrate ($HOBt \cdot H_2O$) (77 mg, 0.50 mmol), and $(i\text{-}Pr)_2NEt$

(170 µL, 1.0 mmol) in DMF (1.8 mL) for 2 h. The first amino acid on the MeDbz linker was manually loaded by Fmoc-D-Tyr(*t*-Bu)-OH (230 mg, 0.50 mmol), 1-[(dimethylamino)(dimethylimino)methyl]-1*H*-[1,2,3]triazolo[4,5-*b*]pyridine 3-oxide hexafluorophosphate (HATU) (190 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 µL, 1.0 mmol) in DMF (1.8 mL) for 2 h. The peptide sequence was constructed by the standard procedure of microwave (MW)-assisted Fmoc-solid phase peptide synthesis (SPPS). After chain assembly, N-terminal amine was protected by $Boc_2O$ (110 mg, 0.50 mmol), (*i*-Pr)$_2$NEt (170 µL, 1.0 mmol) in DMF (1.8 mL) for 2 h. After wash, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (10 mL) for 2 h at 37 °C followed by 0.5 M (*i*-Pr)$_2$NEt in DMF (10 mL) for 30 min at room temperature. This process was repeated three times. Final deprotection and cleavage from the resin were performed by trifluoroacetic acid (TFA)/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in a minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative high-performance liquid chromatography (HPLC) to afford the desired peptide **1** (28 mg, 5.7% yield). MS (ESI): calcd for $C_{154}H_{212}N_{56}O_{39}S$: 3503.80; observed: $[M+5H]^{5+}$ *m/z* = 702.06, $[M+4H]^{4+}$ *m/z* = 877.44, $[M+3H]^{3+}$ *m/z* = 1169.64.

**[Thz$^{27}$]-D-VEGF-A$^{27–57}$ (2).** C-terminal arginine tag was manually loaded on NovaSyn TGR resin (400 mg, 0.10 mmol) by Fmoc-D-Arg(Pbf)-OH (320 mg, 0.50 mmol), OxymaPure (71 mg, 0.50 mmol) and DIC (160 µL, 1.0 mmol) in DMF (1.8 mL) for 2 h. MeDbz linker was then loaded on the resin by Fmoc-MeDbz-OH (190 mg, 0.50 mmol), HBTU (190 mg, 0.50 mmol), HOBt·$H_2O$ (77 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 µL, 1.0 mmol) in DMF (1.8 mL) for 2 h. The first amino acid on the MeDbz linker was manually loaded by Fmoc-D-Arg(Pbf)-OH (320 mg, 0.50 mmol), HATU (190 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 µL, 1.0 mmol) in DMF (1.8 mL) for 2 h. The peptide sequence was constructed by the standard procedure of MW-assisted Fmoc-SPPS. Boc-D-Thz-OH was employed for the coupling of the N-terminal amino acid. After chain assembly, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (10 mL) for 1 h followed by 0.5 M (*i*-Pr)$_2$NEt in DMF (10 mL) for 15 min. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in a minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford the desired peptide **2** (23 mg, 4.9% yield). MS (ESI): calcd for $C_{187}H_{277}N_{45}O_{51}S_3$: 4067.72; observed: $[M+4H]^{4+}$ *m/z* =

1018.46, $[M+3H]^{3+}$ *m/z* = 1357.47.

**D-VEGF-A$^{58–110}$ (3).** By the standard procedure of MW-assisted Fmoc-SPPS described above, the peptide sequence from Lys$^{109}$ to Ile$^{81}$ was constructed from Fmoc-D-Asp(O*t*-Bu)-Wang resin (170 mg, 0.10 mmol). The peptide sequence from Gln$^{80}$ to Cys$^{58}$ was constructed by the standard Fmoc-SPPS. Global deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in a minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford the desired peptide **3** (25 mg, 3.5% yield). MS (ESI): calcd for $C_{246}H_{398}N_{76}O_{81}S_9$: 6004.88; observed: $[M+7H]^{7+}$ *m/z* = 859.28, $[M+6H]^{6+}$ *m/z* = 1002.29, $[M+5H]^{5+}$ *m/z* = 1202.56, $[M+4H]^{4+}$ *m/z* = 1502.74.

**D-VEGF-A$^{1–110}$ (4).** MeNbz-peptide **2** (11 mg, 2.3 μmol) and cysteine-peptide **3** (20 mg, 2.8 μmol) were reacted in ligation buffer (100 mM 4-mercaptophenylacetic acid (MPAA), 50 mM tris(2-carboxyethyl)phosphine (TCEP), 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 1.2 mL) for 1 h at 37 °C. The same amount of a deThz buffer (200 mM methoxyamine·HCl, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 4.0; 1.2 mL) was added and the reaction mixture was incubated for 3 h at 37 °C. After the conversion of the Thz-peptide product to the Cys-peptide, the pH was adjusted to 7.0. Lyophilized peptide **1** (17 mg, 3.5 μmol) was added to the same reaction mixture and incubated for 1 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide **4** (6.3 mg, 18% yield). MS (ESI): calcd for $C_{556}H_{845}N_{163}O_{165}S_{13}$: 12869.63; observed: $[M+14H]^{14+}$ *m/z* = 920.55, $[M+13H]^{13+}$ *m/z* = 991.37, $[M+12H]^{12+}$ *m/z* = 1073.87, $[M+11H]^{11+}$ *m/z* = 1171.45, $[M+10H]^{10+}$ *m/z* = 1288.42, $[M+9H]^{9+}$ *m/z* = 1431.36, $[M+8H]^{8+}$ *m/z* = 1609.95, $[M+7H]^{7+}$ *m/z* = 1840.60.

**Folding of D-VEGF-A (Synthesis of 5).** By the identical procedure of the previous report (1), the folding process of D-VEGF-A was performed. Lyophilized peptide **4** (1.6 mg, 0.10 μmol) was dissolved in 6 M guanidine·HCl solution at the concentration of 20 mg/mL. The solution was diluted 40-fold in a folding buffer (100 mM Tris, 2.0 mM reduced glutathione, 0.4 mM oxidized glutathione, pH 8.4). The solution was incubated for 5 days at room temperature without stirring. The reaction was monitored by analytical HPLC. The folded product was purified by preparative HPLC to afford the desired folded D-VEGF-A **5** (0.65 mg, 20% yield). MS (ESI): calcd for $C_{1112}H_{1674}N_{326}O_{330}S_{26}$: 25723.14; observed: $[M+18H]^{18+}$ *m/z* = 1429.17, $[M+17H]^{17+}$ *m/z* = 1514.05, $[M+16H]^{16+}$ *m/z* =

1608.80, $[M+15H]^{15+}$ *m/z* = 1715.93, $[M+14H]^{14+}$ *m/z* = 1838.23, $[M+13H]^{13+}$ *m/z* = 1979.85.

**D-VHH[B1]$^{1–21}$ (6).** C-terminal arginine tag was manually loaded on Fmoc-NH-SAL resin (440 mg, 0.25 mmol) by Fmoc-D-Arg(Pbf)-OH (810 mg, 1.25 mmol), OxymaPure (180 mg, 1.25 mmol) and DIC (390 µL, 2.5 mmol) in DMF (4.4 mL) for 2 h. Dbz linker was then loaded on the resin by Fmoc-Dbz-OH (470 mg, 1.25 mmol), HBTU (470 mg, 1.25 mmol), HOBt·$H_2O$ (190 mg, 1.25 mmol), and (*i*-Pr)$_2$NEt (430 µL, 2.5 mmol) in DMF (4.4 mL) for 2 h. The first amino acid on the Dbz linker was manually loaded by Fmoc-D-Ser(*t*-Bu)-OH (480 mg, 1.25 mmol), HATU (480 mg, 1.25 mmol), and (*i*-Pr)$_2$NEt (430 µL, 2.5 mmol) in DMF (4.4 mL) for 1 h. The peptide sequence was constructed by the standard procedure of MW-assisted Fmoc-SPPS. After chain assembly, N-terminal amine was protected by $Boc_2O$ (270 mg, 1.25 mmol), (*i*-Pr)$_2$NEt (440 µL, 2.5 mmol) in DMF (4.4 mL) for 2 h. After wash, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (25 mL) for 1 h followed by 0.5 M (*i*-Pr)$_2$NEt in DMF (25 mL) for 15 min. 0.125 mmol resin was used for the next cleavage and purification steps. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in a minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford the desired peptide **6** (56 mg, 16% yield from 0.125 mmol resin). MS (ESI): calcd for $C_{103}H_{169}N_{33}O_{33}$: 2397.68; observed: $[M+4H]^{4+}$ *m/z* = 600.4, $[M+3H]^{3+}$ *m/z* = 800.3, $[M+2H]^{2+}$ *m/z* = 1199.7.

**[Cys(Acm)$^{22}$]-D-VHH[B1]$^{22–49}$ (7).** C-terminal arginine tag was manually loaded on NovaSyn TGR resin (2.0 g, 0.50 mmol) by Fmoc-D-Arg(Pbf)-OH (1.6 g, 2.5 mmol), OxymaPure (360 mg, 2.5 mmol) and DIC (780 µL, 5.0 mmol) in DMF (8.8 mL) for 2 h. MeDbz linker was then loaded on the resin by Fmoc-MeDbz-OH (580 mg, 1.5 mmol), HATU (570 mg, 1.5 mmol), and (*i*-Pr)$_2$NEt (520 µL, 3.0 mmol) in DMF (8.8 mL) for 3 h at 37 °C. The first amino acid on the MeDbz linker was manually loaded by Fmoc-D-Ala-OH (780 mg, 2.5 mmol), HATU (950 mg, 2.5 mmol) and (*i*-Pr)$_2$NEt (870 µL, 5.0 mmol) in DMF (8.8 mL) for 3 h at 37 °C. The peptide sequence was constructed by the standard procedure of MW-assisted Fmoc-SPPS. Boc-D-Cys(Acm)-OH was employed for the coupling of an N-terminal cysteine. After chain assembly, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (50 mL) for 1 h followed by 0.5 M (*i*-Pr)$_2$NEt in DMF (50 mL) for 15 min. Final deprotection and cleavage from the resin

were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in a minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford the desired peptide **7** (330 mg, 16% yield). MS (ESI): calcd for $C_{156}H_{232}N_{50}O_{39}S_2$: 3483.99; observed: $[M+6H]^{6+}$ *m/z* = 581.6, $[M+5H]^{5+}$ *m/z* = 697.7, $[M+4H]^{4+}$ *m/z* = 871.9, $[M+3H]^{3+}$ *m/z* = 1162.3.

**[Cys$^{50}$]D-VHH[B1]$^{50–60}$ (8).** C-terminal arginine tag was manually loaded on NovaSyn TGR resin (2.0 g, 0.50 mmol) by Fmoc-D-Arg(Pbf)-OH (1.6 g, 2.5 mmol), OxymaPure (360 mg, 2.5 mmol) and DIC (780 µL, 5.0 mmol) in DMF (8.8 mL) for 2 h. Dbz linker was then loaded on the resin by Fmoc-(*o*-Boc)Dbz-OH (710 mg, 1.5 mmol), HATU (570 mg, 1.5 mmol), and (*i*-Pr)$_2$NEt (520 µL, 3.0 mmol) in DMF (8.8 mL) for 3 h at 37 °C. The first amino acid on the (*o*-Boc)Dbz linker was manually loaded by Fmoc-D-Tyr(*t*-Bu)-OH (1.1 g, 2.5 mmol), HATU (950 mg, 2.5 mmol) and (*i*-Pr)$_2$NEt (870 µL, 5.0 mmol) in DMF (8.8 mL) for 3 h at 37 °C. The peptide sequence was constructed by the standard procedure of MW-assisted Fmoc-SPPS. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in a minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford desired peptide **8** (240 mg, 27% yield). MS (ESI): calcd for $C_{68}H_{93}N_{19}O_{19}S$: 1512.67; observed: $[M+3H]^{3+}$ *m/z* = 505.2, $[M+2H]^{2+}$ *m/z* = 757.3.

**[Cys$^{61}$]-D-VHH[B1]$^{61–95}$ (9).** C-terminal poly-arginine tag was manually loaded on Fmoc-NH-SAL resin (1.0 g, 0.50 mmol) by Fmoc-D-Arg(Pbf)-OH (1.6 g, 2.5 mmol), OxymaPure (360 mg, 2.5 mmol) and DIC (780 µL, 5.0 mmol) in DMF (8.8 mL) for 2 h each step. Dbz linker was then loaded on the resin by Fmoc-(*o*-Boc)Dbz-OH (710 mg, 1.5 mmol), HATU (570 mg, 1.5 mmol), and (*i*-Pr)$_2$NEt (520 µL, 3.0 mmol) in DMF (8.8 mL) for 3 h at 37 °C. The first amino acid on the (*o*-Boc)Dbz linker was manually loaded by Fmoc-D-Tyr(*t*-Bu)-OH (1.1 g, 2.5 mmol), HATU (950 mg, 2.5 mmol) and (*i*-Pr)$_2$NEt (870 µL, 5.0 mmol) in DMF (8.8 mL) for 3 h at 37 °C. The peptide sequence from Tyr$^{95}$ to Ala$^{75}$ was constructed by the standard procedure of MW-assisted Fmoc-SPPS. The peptide sequence from Asn$^{74}$ to Cys$^{61}$ was constructed by the standard Fmoc-SPPS. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The

resulting precipitate was dissolved in a minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford the desired peptide **9** (400 mg, 13% yield). MS (ESI): calcd for $C_{204}H_{329}N_{69}O_{61}S_2$: 4788.42; observed: $[M+7H+TFA]^{7+}$ *m/z* = 701.3, $[M+6H+TFA]^{6+}$ *m/z* = 817.8, $[M+5H]^{5+}$ *m/z* = 958.7, $[M+4H]^{4+}$ *m/z* = 1198.2, $[M+3H]^{3+}$ *m/z* = 1597.3.

**D-VHH[B1]$^{96–125}$ (10).** By the standard procedure of MW-assisted Fmoc-SPPS, the peptide sequence was constructed from NovaSyn TGR resin (400 mg, 0.10 mmol). Global deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in a minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford desired peptide **10** (14 mg, 3.1% yield). MS (ESI): calcd for $C_{142}H_{210}N_{50}O_{42}S_3$: 3385.73; observed: $[M+5H]^{5+}$ *m/z* = 678.1, $[M+4H]^{4+}$ *m/z* = 847.4, $[M+3H]^{3+}$ *m/z* = 1129.5, $[M+2H]^{2+}$ *m/z* = 1693.9.

**[Cys(Acm)$^{22}$/Cys$^{50}$]-D-VHH[B1]$^{22–60}$ (11).** MeNbz-peptide **7** (440 mg, 100 µmol) and cysteine-peptide **8** (180 mg, 100 µmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 51 mL) for 2 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide **11** (420 mg, 77% yield). MS (ESI): calcd for $C_{208}H_{304}N_{62}O_{55}S_3$: 4649.28; observed: $[M+6H]^{6+}$ *m/z* = 775.7, $[M+5H]^{5+}$ *m/z* = 930.8, $[M+4H]^{4+}$ *m/z* = 1163.2, $[M+3H]^{3+}$ *m/z* = 1550.6.

**D-VHH[B1]$^{22–95}$ (12).** An $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 1.85 mL) was added to a solution of Dbz-peptide **11** (400 mg, 74 µmol) in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 19 mL) at -20 °C. After the mixture was stirred for 30 min, MPAA buffer (400 mM MPAA buffer, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 19 mL) was added to the reaction mixture. The mixture was warmed to room temperature, and the pH was adjusted to 7.0. Lyophilized peptide **9** (270 mg, 44 µmol) was added to the mixture and incubated for 1 h at 37 °C. The reaction was monitored by analytical HPLC. After the reaction was completed, unreacted reagents were separated by preparative HPLC and the desired peptide (with small amount of impurities) was lyophilized. Lyophilized peptide mixture (480 mg) was reacted in desulfurization buffer (20 mM 2,2'-azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride] (VA-044), 100 mM 2-mercaptethanesulfonate (MESNa), 250 mM TCEP, 6 M guanidine·HCl, 200 mM

phosphate buffer, pH 6.5; 88 mL) for 2 h at 37 °C. The reaction was monitored by LC-MS. After the reaction was completed, unreacted reagents were separated by preparative HPLC and the desired peptide (with small amount of impurities) was lyophilized. Lyophilized peptide mixture (420 mg) and $PdCl_2$ (69 mg, 390 μmol) were reacted in 200 mM phosphate buffer containing 6 M guanidine·HCl (20 mL) for 30 min at 37 °C and treated with small amounts of DTT. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford desired peptide **12** (53 mg, 6.7% yield). MS (ESI): calcd for $C_{396}H_{607}N_{123}O_{113}S_3$: 8995.14; observed: $[M+13H+TFA]^{13+}$ *m/z* = 701.8, $[M+12H+TFA]^{12+}$ *m/z* = 760.0, $[M+11H+TFA]^{11+}$ *m/z* = 829.1, $[M+10H+TFA]^{10+}$ *m/z* = 911.9, $[M+9H+TFA]^{9+}$ *m/z* = 1013.1, $[M+8H]^{8+}$ *m/z* = 1125.5, $[M+7H]^{7+}$ *m/z* = 1286.2, $[M+6H]^{6+}$ *m/z* = 1500.3, $[M+5H]^{5+}$ *m/z* = 1800.2.

**D-VHH[B1]$^{1–95}$ (13).** Nbz-peptide **6** (17 mg, 6.1 μmol) and cysteine-peptide **12** (32 mg, 3.0 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 1.5 mL) for 1 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide **13** (25 mg, 63% yield). MS (ESI): calcd for $C_{485}H_{757}N_{149}O_{143}S_3$: 11059.47; observed: $[M+16H+TFA]^{16+}$ *m/z* = 699.3, $[M+14H+TFA]^{14+}$ *m/z* = 799.0, $[M+13H+TFA]^{13+}$ *m/z* = 860.4, $[M+12H+TFA]^{12+}$ *m/z* = 932.1, $[M+11H+TFA]^{11+}$ *m/z* = 1016.9, $[M+10H]^{10+}$ *m/z* = 1107.0, $[M+9H]^{9+}$ *m/z* = 1229.8, $[M+8H]^{8+}$ *m/z* = 1383.4, $[M+7H]^{7+}$ *m/z* = 1580.9.

**D-VHH[B1]$^{1–125}$ (14).** An $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 44 μL) was added to a solution of Dbz-peptide **13** (23 mg, 1.8 μmol) in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 440 μL) at -20 °C, the reaction was continued for 30 min. An MPAA buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 440 μL) was then added to the reaction, and the reaction was continued for 1 min. The reaction was monitored by analytical HPLC. After the reaction was completed, unreacted reagents were separated by preparative HPLC and the desired peptide (with small amount of impurities) was lyophilized. Lyophilized thioester (14 mg) and cysteine-peptide **10** (6.4 mg, 1.5 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 1.8 mL) for 2 h at 37 °C. The reaction was monitored by analytical HPLC. The crude products were purified by preparative HPLC to afford the desired peptide **14** (11 mg, 37% yield). MS (ESI): calcd for $C_{596}H_{910}N_{180}O_{180}S_6$: 13669.27; observed: $[M+16H]^{16+}$ *m/z* = 855.2, $[M+15H]^{15+}$ *m/z* = 912.2, $[M+14H]^{14+}$ *m/z* = 977.4, $[M+13H]^{13+}$ *m/z* = 1052.6, $[M+12H]^{12+}$ *m/z* = 1140.0,

$[M+11H]^{11+}$ *m/z* = 1243.6, $[M+10H]^{10+}$ *m/z* = 1367.9, $[M+9H]^{9+}$ *m/z* = 1520.0, $[M+8H]^{8+}$ *m/z* = 1709.8.

**Folding of D-VHH[B1] (Synthesis of 15).** Lyophilized peptide **14** was dissolved in 20 mM phosphate buffer (pH 7.5) containing 6 M guanidine·HCl at a concentration of 0.25 mg/mL in a dialysis cassette (Molecular Weight Cut-Off = 10,000) and dialyzed against 500 volumes of a series of dialysis buffers (100 mM Tris, 5 mM reduced glutathione; 0.5 mM oxidized glutathione, pH 8.4) containing 4 M, 2 M, 1 M, 0.5 M guanidine·HCl at 4 °C for 6 h each. Subsequently, the folded protein was dialyzed against 500 volumes of a series of PBS (pH 7.4) containing 0.25 M, 0.1 M, and 0 M guanidine·HCl at 4 °C for 6 hours each. Finally, the folded protein in 0 M guanidine·HCl was dialyzed against PBS (pH 7.4) twice at 4 °C. The solution was filtered and concentrated by ultrafiltration using an MWCO 10,000 centrifugal filtration membrane (Millipore, Amicon-Ultra 10 kDa).

**Preparation of Recombinant VHH.** The VHH genes from the selected clones were re-cloned into the pGEX6p1 vector and transformed into the competent SHuffle5403 stain. The bacteria were then cultured in LB medium supplemented with ampicillin (50 μg/mL) overnight at 37 °C. Protein expression was induced with 100 μM IPTG overnight at 20 °C. Cells were pelleted by centrifugation (3,600 × *g*, 20 min, 4 °C) and resuspended in PBS containing protease inhibitor (Nacalai Tesque, Inc.). After sonication and the addition of Triton X-100, the cell lysates were centrifuged at 14,200 × *g* for 20 min at 4 °C. The supernatants were incubated with Glutathione-Sepharose 4B (GE Healthcare) overnight at 4 °C. The glutathione S-transferase (GST) tags were then removed by incubation with PreScission protease (GE Healthcare) in cleavage buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.01% Triton X-100) overnight at 4 °C. VHH[E1] was further purified by gel filtration chromatography (TSKgel G2000$SW_{XL}$, TOSOH) to be used in the immunogenicity test. The purity of all proteins was confirmed by SDS-PAGE.

**SPR Analysis.** The binding affinities of the VHHs to L-/D-VEGF-A were measured by SPR (Biacore X100, Cytiva, Japan). During the procedure, HBS-EP buffer (Cytiva) was used as the running buffer (at 25 °C) and 1 M NaCl solution was used as the regeneration solution. D-VEGF-A was immobilized on a Sensor Chip CM5 (534.1 RU) using an amine coupling kit (EDC/NHS). L-VEGF-A (Biotinylated human VEGF-A121, Acro Biosystems, cat# VE1-H82E7) was immobilized on a Sensor Chip SA (258.8 RU). All analytes were evaluated during a 2-min contact period, followed by a 4-min dissociation period at a flow rate of 30 μL/min. $K_D$ values were calculated by steady-state binding analysis of data generated from experiments performed in triplicate.

**CD Spectra.** Synthetic D-VHH[B1] and recombinant L-VHH[B1] were diluted in PBS (pH7.4) so that the concentration of each protein was 10 μM. CD spectra of each protein were recorded using a circular dichroism spectrometer (JACSO J-1500) at 20 °C.

**Immunogenicity Test.** The animal care and experimental procedures used in this study were approved by the Animal Care Committee of Kyoto University. Five-week-old female BALB/c mice were purchased from Japan SLC, Inc. (Hamamatsu, Japan). Five BALB/c mice (female, 5 weeks old) per group were given intraperitoneal injections on day 0 with 25 μg (per mouse) of recombinant L-VHH[E1] or synthetic D-VHH[B1] antigen emulsified in Freund's Complete Adjuvant (Wako). On days 14 and 28, the mice were further injected with 25 μg (per mouse) of recombinant L-VHH[E1] or synthetic D-VHH[B1] antigen emulsified in Freund's Incomplete Adjuvant (Wako). Sera were collected on days 0, 14, 28, and 42 and analyzed by ELISA. PBS (pH 7.4) containing 0.025% Tween 20 was used for all ELISA washing and dilution steps, unless otherwise stated. Briefly, MaxiSorp plates (Thermo Scientific) were coated overnight at 4 °C with either recombinant L-VHH[E1] or synthetic D-VHH[B1] in 50 mM sodium carbonate buffer (pH 9.4, 50 μL/well; 1.0–1,000 ng/mL). After coating, the wells were washed with PBS three times and blocked with PBS containing 3% BSA (150 μL/well) for 2 h. After three washes, the serum samples (diluted 1:1,000) from each mouse (50 μL/well; collected on days 0, 14, 28, or 42) were applied to the plates and incubated for 1 h. After three washes, the anti-mouse IgG (H+L)-HRP conjugate (Jackson) (diluted 1:5,000, 50 μL/well) was added to the wells and incubated for 30 min. After four washes, the wells were incubated for 15 min with TMB solution and the reaction was stopped with 1 M $H_2SO_4$ (50 μL/well). Absorbance was measured at 450 nm using a microplate absorbance reader (infinite 200, TECAN).

References


1. Hanke, L.; Sheward, D. J. Pankow, A., Vidakovics, L. P.; Karl, V.; Kim, C.; Urgard, E.; Smith, N. L.; Astorga-Wells, J.; Ekström, S.; Coquet, J. M.; Mclnerney, G. M.; Murrell, B. *Sci. Adv.* **2022**, *8*, eabm0220.
2. Sabir, J. S. M.; Atef, A.; El-Domyati, F. M.; Edris, S.; Hajrah, N.; Alzohairy, A. M.; Bahieldin, A. *C. R. Biologies* **2014**, *337*, 244–249.
3. Doshi, R.; Chen, B. R.; Vibat, C. R. T.; Huang, N.; Lee, C.-W.; Chang, G. *Sci. Rep.* **2014**, *4*, 6760.
4. McMahon, C.; Baier, A. S.; Pascolutti, R.; Wegrecki, M.; Zheng, S.; Ong, J. X.; Erlandson, S. C.; Hilger, D.; Rasmussen, S. G. F.; Ring, A. M.; Manglik, A.; Kruse, A. C. *Nat. Struct. Mol. Biol.* **2018**, *25*, 289–296.
5. Liu, B.; Yang, D. *Int. J. Mol. Sci.* **2022**, *23*, 1482.
6. Nguyen, V. K.; Hamers, R.; Wyns, L.; Muyldermans, S. *EMBO J.* **2020**, *19*, 921–930.
7. Mueller, L. K.; Baumruck, A. C.; Zhdanova, H.; Tietze, A. A. *Front. Bioeng. Biotechnol.* **2020**, *8*, 162.
8. Aoki, K.; Manabe, A.; Kimura, H.; Katoh, Y.; Inuki, S.; Ohno, H.; Nonaka, M.; Oishi, S. *Bioconjug. Chem.* **2023**, *34*, 2055–2065.
9. Duggan, S. *Drugs* **2018**, *78*, 1639–1642.
10. Ferrara, N.; Mass, R. D.; Campa, C.; Kim, R. *Annu. Rev. Med.* **2007**, *58*, 491–504.
11. Yue, H.; Li, Y.; Yang, M.; Mao, C. *Adv. Sci.* **2022**, *9*, 2103645.
12. Higashi, K.; Oda, S.; Fujii, M.; Nishida, F.; Matsumoto, H.; Morise, J.; Oka, S.; Nonaka, M. *J. Biochem.* **2024**, *175*, 85–93.
13. Wilton, E. E.; Opyr, M. P.; Kailasam, S.; Kothe, R. F.; Wieden, H.-J. *ACS Synth. Biol.* **2018**, *7*, 2480–2484.
14. Shimko, T. C.; Fordyce, P. M.; Orenstein, Y. *Bioinfomatics* **2020**, *36*, 3357–3364.
15. Lee, C. V.; Liang, W.-C.; Dennis, M. S.; Eigenbrot, C.; Sidhu, S. S.; Fuh, G. *J. Mol. Biol.* **2004**, *340*, 1073–1093.
16. Garamella, J.; Marshall, R.; Rustad, M.; Noireaux, V. *ACS Synth. Biol.* **2016**, *5*, 344–355.
17. Pielou, E. C. *J. Theoret. Biol.* **1966**, *13*, 131–144.
18. Mandal, K.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2011**, *50*, 8029–8033.
19. Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc.* **2015**, *137*, 7197–7209.
20. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.

21. Hochuli, E.; Bannwarth, W.; Dobeli, H.; Gentz, R.; Stuber, D. *Nat. Biotechnol.* **1988**, *6*, 1321–1325.
22. Yamauchi, O.; Odani, A. *Pure Appl. Chem.* **1996**, *68*, 469–496.
23. Wang, J.-X.; Fang, G.-M.; He, Y.; Qu, D.-L.; Yu, M.; Hong, Z.-Y.; Liu, L. *Angew. Chem., Int. Ed.* **2015**, *54*, 2194–2198.
24. Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248–9252.
25. Maity, S. K.; Jbara, M.; Laps, S.; Brik, A. *Angew. Chem., Int. Ed.* **2016**, *55*, 8108–8112.
26. Collins, J. M.; Porter, K. A.; Singh, S. K.; Vanier, G. S. *Org. Lett.* **2014**, *16*, 940–943.

# Chapter 2. Synthetic Studies of Therapeutic Targets for Mirror-Image Screening

## Section 1. Synthetic Study of Full-Length Hepatitis B Virus Core Protein and Its Capsid Assembly

*Summary: Chronic infection with hepatitis B virus (HBV) is a major cause of cirrhosis and liver cancer. Capsid assembly modulators can induce error-prone assembly of HBV core proteins to prevent the formation of infectious virions, representing promising candidates for treating chronic HBV infections. To explore novel capsid assembly modulators from unexplored mirror-image libraries, the author investigated the synthetic process of the HBV core protein for preparing the mirror-image target protein. In this Section, the chemical synthesis of full-length HBV core protein [Cp183(C183A)] containing an arginine-rich nucleic acid-binding domain at the C-terminus is presented. Sequential ligations using four peptide segments enabled the synthesis of Cp183(C183A) via convergent and C-to-N direction approaches. After refolding under appropriate conditions, followed by the addition of nucleic acid, the synthetic Cp183(C183A) assembled into capsid-like particles.*

Chronic infection with hepatitis B virus (HBV) is a leading cause of cirrhosis and hepatocellular carcinoma.[1] Although an effective vaccine is available, approximately 290 million people worldwide have chronic hepatitis B, and more than 800,000 people die annually from HBV-related diseases, as estimated by the World Health Organization.[2] Several drugs have been approved for treating chronic HBV infections, including nucleoside and nucleotide analogs targeting the viral reverse transcriptase and interferon-α to stimulate host immunity.[1] These agents are effective in inhibiting HBV replication and suppressing hepatitis; however, the treatment is not curative because of their inability to eliminate covalently closed circular DNA (cccDNA), which is the reservoir for persistent HBV infection.[3] Therefore, developing novel therapeutic agents via an alternative mechanism(s) is needed.[4]

The HBV core protein, a building block of nucleocapsids, plays essential roles during multiple stages of the viral replication cycle, including nucleocapsid assembly, encapsidation of the pregenomic RNA (pgRNA), reverse transcription, and DNA synthesis.[5,6] Many compounds targeting the HBV core protein have been identified.[7–10] These compounds are classified into two categories based on the post-treatment capsid states. Class I compounds (e.g., heteroaryldihydropyrimidines) induce misfolding of the

core protein structure,[11] whereas class II compounds (e.g., phenylpropenamides and sulfamoylbenzamides) accelerate the formation of empty capsids without pgRNA.[12] Both of these capsid assembly modulators (CAMs) promote error-prone assembly of core proteins, leading to the prevention of infectious virion formation. Thus, the HBV core protein is a promising target for developing anti-HBV agents.[13] Some CAMs targeting the HBV core protein are currently in phase 1 and 2 clinical trials for treating HBV infection.[4]

To identify novel drug candidates to treat chronic HBV infection by mirror-image screening, the author focused on the HBV core protein as the target protein. Developing synthetic and folding protocols of the core protein to obtain functional protein is essential for mirror-image screening because mirror-image proteins cannot be prepared by recombinant technology.[14] The full-length HBV core protein (Cp183) consists of an assembly domain (Cp149: Met[1]–Val[149]) for capsid formation and a C-terminal arginine-rich nucleic acid-binding domain (Arg[150]–Cys[183]) (Figure 1).[15,16] Cp149 was reported to be expressed in *Escherichia coli* and spontaneously assemble to form capsid particles.[17] The Cp149 capsids disassemble into core protein dimers in the presence of guanidine or urea, whereas the reassembly of empty Cp149 capsids was triggered by increasing the ionic strength.[18,19] Previously, Tsuda *et al.* reported the chemical synthesis of the C-terminal truncated HBV core protein (Cp149) (Figure 2).[20] However, unlike recombinantly expressed Cp149, the preparation of capsid particles from the synthetic Cp149 failed. The author postulated that nucleic acid association with the C-terminal nucleic acid-binding domain promotes *in vitro* assembly to form capsid particles from synthetic proteins.[21] Based on these previous findings, the author selected the full-length core protein (Cp183) as an alternative synthetic target. In this study, the author describes the synthetic process and *in vitro* capsid assembly of the full-length HBV core protein.

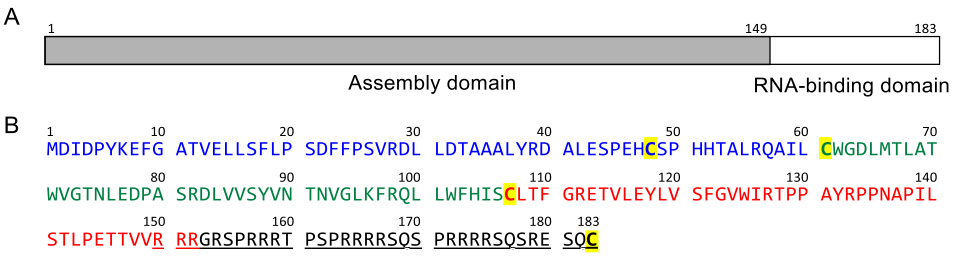


**Figure 1.** Domain structure and sequence of the HBV core protein (Cp183). (A) Schematic representation of the HBV core protein domain architecture, showing the assembly domain (amino acids 1–149) and the nucleic acid-binding domain (amino acids 150–183). (B) Sequence of the HBV core protein (subtype adyw).[15] Cysteine residues are highlighted in yellow. The underlined sequence is the nucleic acid-binding domain.

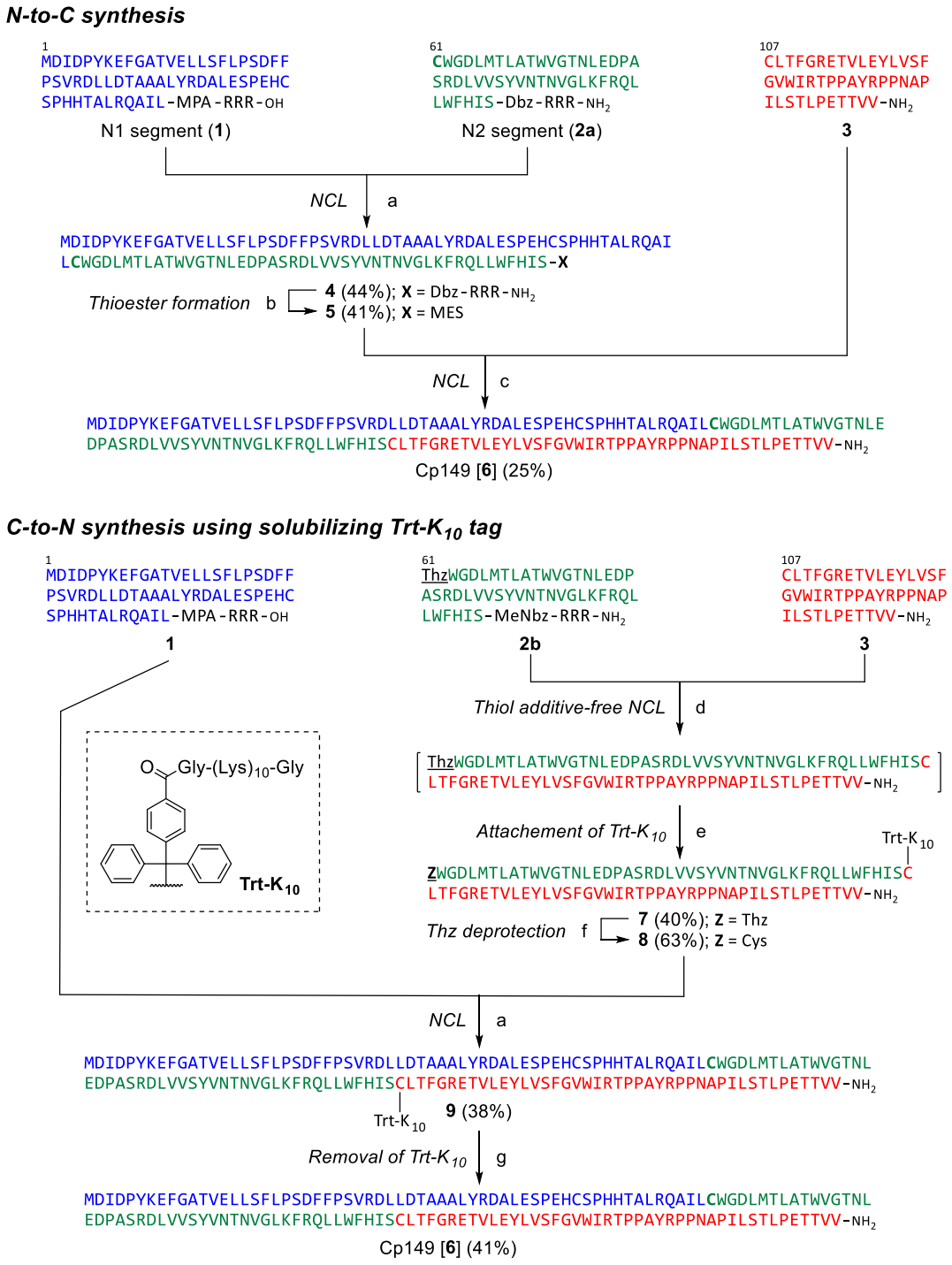


**Figure 2.** Previous synthetic studies on Cp149.[20] *Reagents and conditions*: (a) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (b) $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, then MESNa and TCEP; (c) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0) containing 25% NMP; (d) 1,2,4-triazole, TCEP, 6 M guanidine·HCl, and 100 mM phosphate buffer (pH 7.1); (e) Trt(OH)-$K_{10}$ and TFA; (f) methoxyamine, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 4.0); (g) TFA/TIS (95:5).

**Design of the Synthetic Process for Cp183.** Cp183 contains four cysteine residues ($Cys^{48}$, $Cys^{61}$, $Cys^{107}$, and $Cys^{183}$) in the 183-residue sequence (Figure 1). Among these, N-proximal cysteines ($Cys^{48}$, $Cys^{61}$, and $Cys^{107}$) in the assembly domain form intermolecular disulfide bonds to yield Cp183 dimers,[22] whereas $Cys^{183}$ in the nucleic acid-binding domain is not required for capsid assembly.[23] Thus, the author substituted $Cys^{183}$ with Ala to avoid the potential side reaction(s), including epimerization[24] and β-piperidinyl alanine formation[25] at this Cys residue during peptide synthesis. Additionally, in the previous synthetic study of Cp149, attachment of a trityl-based solubilizing tag (Trt-$K_{10}$) to a Cys residue was required because the poor solubility of the peptide segments hampered the ligation and purification processes (Figure 2).[20] The author hypothesized that the C-terminal arginine-rich domain in the Cp183 sequence should enhance the solubility of the peptide segments, which works as a solubilizing tag like the previously developed ones.[26,27]

Several strategies have also been developed to extend the scope of native chemical ligation (NCL)[28,29] to various ligation junctions beyond Cys or Ala[30]; however, ligation auxiliaries should be prepared in advance for practical applications.[31,32] Thus, the author planned to use the Ag-free thioester method[33] because the C-terminal sequence ($Arg^{150}$–$Ala^{183}$) of Cp183(C183A) has no Cys or Ala residues in appropriate positions for NCL. This method allows direct amide bond formation between a thioester and amine in the presence of protecting groups for other amino and thiol groups. However, the author posited that the ligation reaction proceeds selectively between unprotected peptide segments in this case because the C-terminal sequence of Cp183(C183A) contains no Lys and Cys residues. The author selected $Gly^{153}$–$Arg^{154}$ as the ligation site to prevent possible epimerization during ligation.

**Synthesis of Cp183(C183A) by the Convergent Route.** Initially, the author planned to synthesize Cp183(C183A) by a convergent ligation strategy using the intermediate segments for Cp149 synthesis (Figure 3).[20] According to the previous synthetic study of Cp149,[20] the author speculated that N-terminal segment **5** would be employed for the synthesis of Cp183(C183A). Two peptide segments were newly designed for the synthesis of Cp183(C183A): the C1 segment **10** ([$Thz^{107}$]-$Cp183^{107–153}$) with a thiazolidine carboxylic acid (Thz) at the N-terminus for temporary protection of Cys and an aryl thioester at the C-terminus; and the C2 segment **11** ([$Ala^{183}$]-$Cp183^{154–183}$). For the preparation of C1 segment **10**, the protected peptide resin was synthesized with an *o*-amino(methyl)aniline (MeDbz)[35] moiety, which was converted to the *N*-acyl-*N′*-methylacylurea (MeNbz) form by treatment with 4-nitrophenyl chloroformate and (*i*-

$Pr)_2$NEt. Global deprotection and cleavage from the resin followed by reaction with 4-mercaptophenylacetic acid (MPAA) provided peptide thioester **10**. Similarly, the C2 segment **11** was obtained using the Fmoc-Ala-Wang resin in a satisfactory yield.

With the designed peptide segments in hand, the author investigated the sequential ligations. The MPAA thioester **10** and peptide **11** were dissolved in DMSO containing 3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine (HOOBt) and ($i$-Pr)$_2$NEt.[33] The coupling proceeded efficiently to provide thiazolidine-protected **12**. Subsequent methoxyamine-mediated deprotection of thiazolidine[36] afforded the desired C-terminal half intermediate **13**. The final ligation between thioester **5** and peptide **13** proceeded to afford the desired full-length Cp183(C183A) **14** in 9.0% yield. In this route, the author successfully constructed the full-length sequence of Cp183(C183A); however, the low ligation efficacy and poor solubility of the hydrophobic intermediate **5** were unsuitable for scale-up synthesis.

**Synthesis of Cp183(C183A) by C-to-N Route.** Next, the author investigated the synthesis of Cp183(C183A) by a C-to-N ligation strategy to avoid using hydrophobic intermediate **5** (Figure 4). The author hypothesized that the C-terminal arginine-rich domain would enhance the solubility of all intermediate peptide segments. Because the N1 segment **1** and C-terminal half **13** can be used for sequential ligations via the C-to-N

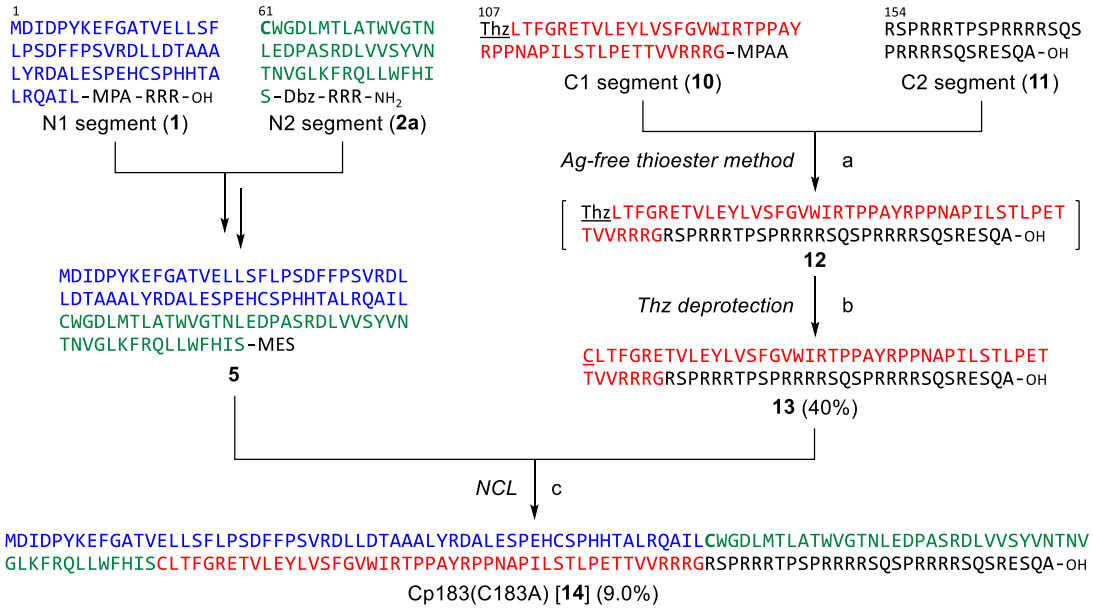


**Figure 3.** Synthesis of Cp183(C183A) via the convergent approach. *Reagents and conditions*: (a) ($i$-Pr)$_2$NEt, HOOBt, and DMSO; (b) methoxyamine, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 4.0); (c) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0).

direction, only the N2 segment **2b** ([Thz$^{61}$]-Cp183$^{61–106}$) was newly synthesized, which contains a Thz for N-terminal temporary protection and a MeNbz linker for C-terminal activation.[20] The resulting peptides **2b** and **13** were subjected to the MPAA-mediated NCL condition. The formation of Thz-protected intermediate **15** was observed when the starting material **2b** was mostly consumed after a 3-hour incubation. The subsequent methoxyamine-mediated deprotection provided the Cys-free intermediate **16** in 24% overall yield (two steps from the peptide segment **2b**). The final ligation between thioester **1** and peptide **16** afforded the desired full-length Cp183(C183A) **14** in sufficient yield (32%). In this C-to-N route, all ligation and purification processes proceeded more smoothly than in the convergent route, probably because all intermediate peptides (i.e., **13** and **16**) contained the arginine-rich sequence. Indeed, these peptides were more hydrophilic than peptide **5** used for the convergent route in HPLC analysis.

**Folding and Capsid Assembly of Synthetic Cp183(C183A).** With full-length Cp183(C183A) in hand, the author investigated the folding conditions to obtain the synthetic protein for capsid assembly. The predominant HBV capsid consists of 120 copies of the core protein dimer arranged in a $T = 4$ icosahedral symmetry.[37] Zlotnick *et al.* reported that Cp183 expressed in *E. coli* assembles to form empty capsids or RNA-filled capsids *in vitro* under appropriate conditions, including ionic strength, Cp183 dimer

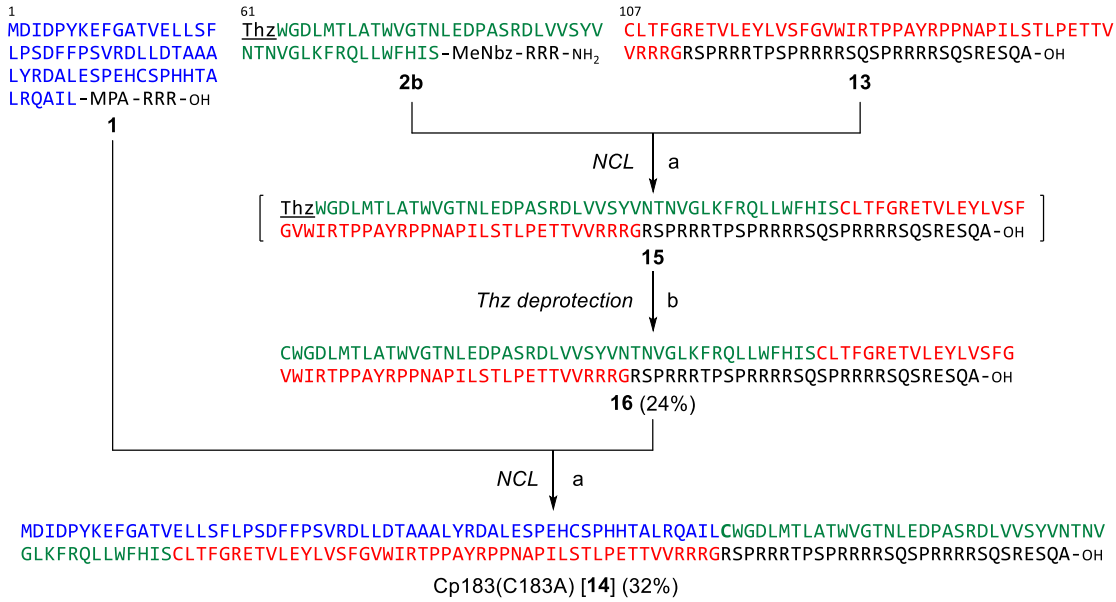


**Figure 4.** Synthesis of Cp183(C183A) via C-to-N native chemical ligations. *Reagents and conditions*: (a) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (b) methoxyamine, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 4.0).

concentration and the presence of nucleic acids.[21] In contrast to the high stability of Cp149 even under low ionic strength conditions, the Cp183 dimer precipitated when the guanidine concentration was less than 0.5 M. According to the reported procedure for capsid assembly from recombinant Cp183,[21] the author initially investigated the assembly conditions from synthetic Cp183 for empty capsids. After lyophilized synthetic Cp183(C183A) was denatured in 6 M guanidine, the concentration of guanidine was gradually reduced using dialysis steps to 1.5 M. Circular dichroism (CD) analysis of Cp183(C183A) in 1.5 M guanidine exhibited the negative band at 222 nm indicative of the existence of an α-helix structure, which is consistent with the canonical secondary structure of the Cp183 monomer (Figure 5A).[38] The negative minimum at 222 nm disappeared when the concentration of guanidine was reduced further from 1.5 M. The negative band at 222 nm was not observed in 1.5 M NaCl (Figure 5B). These results revealed that 1.5 M guanidine used in a previous report is optimal for stabilizing the Cp183 capsid dimer.[21] The author investigated several other conditions to prepare empty capsids in low-ionic-strength buffer, including the concentration of Cp183(C183A), the type of denaturing agents and the use of additives or molecular chaperons;[38,39] however, synthetic Cp183(C183A) was highly prone to aggregation.

The author next investigated the *in vitro* assembly of nucleic acid-filled capsids, which contain single-strand DNA (ssDNA) in the particles (Figure 6A). After preparing the Cp183(C183A) dimer in 1.5 M guanidine using the abovementioned protocol, ssDNA was added to this protein solution. Subsequently, the mixture was diluted 3-fold (to 0.5 M guanidine) according to the assembly protocol for recombinant Cp183.[21] Particle formation was observed by negative-stain electron microscopy (Figure 6B). The size of

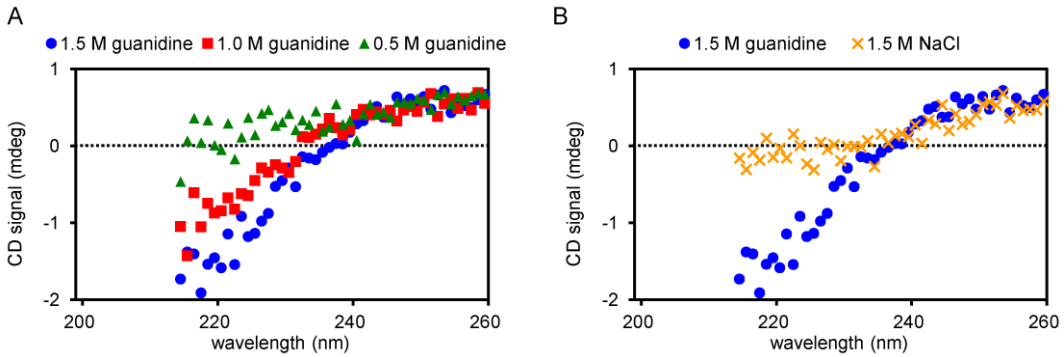


**Figure 5.** CD analysis of the synthetic HBV core protein under various conditions. (A) CD spectra of synthetic Cp183(C183A) at various concentrations of guanidine. (B) Comparison of synthetic Cp183(C183A) CD spectra in a 1.5M guanidine or NaCl solution.

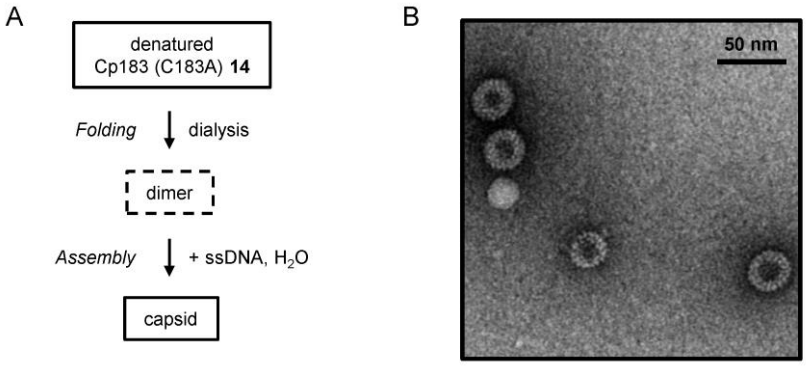


**Figure 6.** Assembly of capsids from synthetic Cp183(C183A). (A) Process for folding and assembly of capsids. (B) Negative-stain electron microscope image of viral capsid-like particles from synthetic Cp183(C183A). *Conditions*: 1.0 μM peptide **14** [synthetic Cp183(C183A)], 13 μM ssDNA, 0.5 M guanidine·HCl, 17 mM HEPES, 0.67 mM DTT, pH 7.5.

the observed particles was 30–40 nm, which is consistent with the reported $T = 3$ (~32 nm) or $T = 4$ (~36 nm) HBV capsids.[21] Notably, a high concentration of synthetic Cp183 for the assembly experiment resulted in aggregation without forming the expected particles. Although further optimization of assembly conditions in the presence of ssDNA is required, these results demonstrated that the synthetic full-length HBV core protein may be used for preparing homogenous samples for mirror-image screening. To the author's knowledge, this is the first report describing the preparation of HBV capsids from synthetic proteins.

In this study, the author established a synthetic process for preparing the full-length HBV core protein [Cp183(C183A)] using NCL and thioester methods. Full-length Cp183(C183A) was constructed by sequential ligations from four peptide segments by a convergent or C-to-N route. For synthesis using the C-to-N direction, the C-terminal arginine-rich domain improved the solubility of all intermediate peptides. The resulting synthetic Cp183(C183A) was appropriately refolded to form HBV capsid particles in the presence of ssDNA. The established procedures of synthetic Cp183(C183A) would be applicable to the preparation of mirror-image Cp183 (D-Cp183). Mirror-image screening using D-Cp183 from natural product libraries having unique, complex, and $sp^3$-carbon-rich scaffolds[40] as well as antibody-like protein libraries such as VHHs[41] would provide novel therapeutic and diagnostic agents for HBV infection (Figure 7).

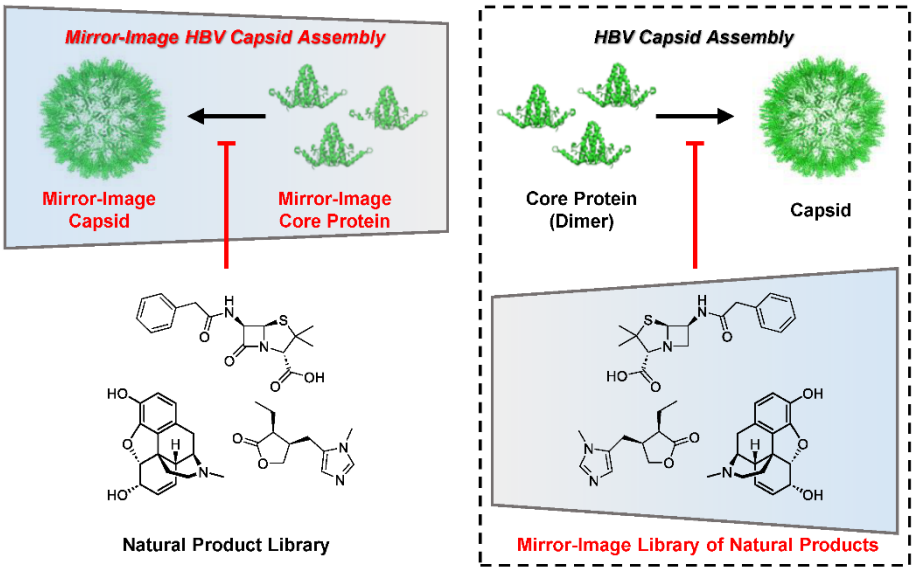


**Figure 7.** Screening concept for HBV capsid assembly modulators from a virtual mirror-image library of natural products. Screening of the natural product library using the mirror-image HBV core protein corresponds to the mirror-image library of natural products using the natural HBV core protein in a mirror. The HBV core protein structure was obtained from the Protein Data Bank (PDB ID: 6HTX).

Experimental Section

**General Procedure of Peptide Synthesis.** All reagents and solvents were purchased from Watanabe Chemical Industries, Ltd. (Hiroshima, Japan), Kokusan Chemical Industries, Ltd. (Kanagawa, Japan), Sigma-Aldrich JAPAN (Tokyo, Japan), Wako Pure Chemical Industries, Ltd. (Osaka, Japan), Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan) or Nacalai Tesque, Inc. (Kyoto, Japan). For analytical high-performance liquid chromatography (HPLC), a Cosmosil 5C18-AR300 column (4.6 × 250 mm, Nacalai Tesque, Inc.), Cosmosil 5C4-AR300 column (4.6 × 150 mm or 4.6 × 250 mm, Nacalai Tesque, Inc.), or DAISOPAC SP-120-5-ODS-BIO (Osaka Soda, 4.6 × 150 mm) was employed with a linear gradient of $CH_3CN$ containing 0.05% or 0.1% (v/v) TFA at a flow rate of 1 mL/min (25 °C or 40 °C). The products were detected by UV absorbance at 220 nm. For preparative HPLC, a Cosmosil 5C18-AR300 column (20 × 250 mm, Nacalai Tesque, Inc.) or a Cosmosil 5C4-AR300 column (20 × 150 mm, Nacalai Tesque, Inc.) was employed with a linear gradient of $CH_3CN$ containing 0.05% or 0.1% TFA at a flow rate of 8 mL/min (room temperature). All peptides were characterized by ESI-MS (micromass ZQ, Waters). Peptide **2b** was synthesized according to the previously reported procedure.[20] Peptides **1** and **5** were provided by Peptide Institute, Inc. (Osaka, Japan).[20]

**Solid-Phase Peptide Synthesis.** Standard Fmoc-based solid-phase peptide synthesis (Fmoc-SPPS) was performed using an automatic peptide synthesizer (PSSM-8, Shimadzu, Japan). The following side chain protected amino acids were employed: Arg(Pbf), Asn(Trt), Asp(O*t*-Bu), Cys(Trt), Gln(Trt), Glu(O*t*-Bu), His(Trt), Lys(Boc), Ser($^t$Bu), Thr(*t*-Bu), and Tyr(*t*-Bu). Fmoc protected amino acids (5 equiv.) were coupled using HBTU (5 equiv.), HOBt·$H_2O$ (5 equiv.) and (*i*-Pr)$_2$NEt (10 equiv.) in DMF for 60 min twice. Fmoc protection group was deprotected by 20% piperidine in DMF for 4 min twice unless otherwise stated.

**[Thz$^{107}$]-Cp183$^{107\text{-}153}$-MPAA (10).** By the standard procedure of automated Fmoc-SPPS, the peptide sequence was constructed from Fmoc-Gly-MeDbz-[Arg(Pbf)]$_3$-Rink amide resin (0.16 mmol). For the coupling of an N-terminal cysteine in peptide **3**, Boc-Thz-OH was employed with DIC/Oxyma pure in DMF for 2 h. Then, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (16 mL) for 1 h followed by 0.5 M (*i*-Pr)$_2$NEt in DMF (16 mL) for 15 min. Global deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of MPAA buffer

(100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 100 mM phosphate buffer, pH 7.2) and incubated for 1 h. The crude products were purified by preparative HPLC to afford desired peptide **3** (41 mg, 4.7% yield). MS (ESI): calcd for $C_{253}H_{393}N_{67}O_{66}S_2$: 5493.45; observed: $[M+6H]^{6+}$ *m/z* = 916.85, $[M+5H]^{5+}$ *m/z* = 1099.99, $[M+4H]^{4+}$ *m/z* = 1374.55, $[M+3H]^{3+}$ *m/z* = 1831.68.

**[Ala$^{183}$]-Cp183$^{154–183}$ (11).** By the standard procedure of automated Fmoc-SPPS, the peptide sequence was constructed from Fmoc-Ala-Wang resin (0.025 mmol). Global deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The crude products were purified by preparative HPLC to afford desired peptide **11** (41 mg, 38% yield). MS (ESI): calcd for $C_{146}H_{264}N_{72}O_{44}$: 3732.18; observed: $[M+6H]^{6+}$ *m/z* = 622.99, $[M+5H]^{5+}$ *m/z* = 747.13, $[M+4H]^{4+}$ *m/z* = 933.93, $[M+3H]^{3+}$ *m/z* = 1245.15, $[M+2H]^{2+}$ *m/z* = 1866.85.

**[Ala$^{183}$]-Cp183$^{107–183}$ (13).** MPAA thioester **10** (10 mg, 1.8 µmol) and peptide **11** (7.7 mg, 2.0 µmol) were reacted with HOOBt (8.9 mg, 55 µmol) and (*i*-Pr)$_2$NEt (6.2 µL, 36 µmol) in DMSO (600 µL) for 15 h at room temperature. Then, deThz buffer (1 M methoxyamine hydrochloride, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 4.0; 3.4 mL) was added and the reaction mixture was incubated for 3 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide **13** (6.5 mg, 40% yield). MS (ESI): calcd for $C_{390}H_{649}N_{139}O_{108}S$: 9045.41; observed: $[M+8H]^{8+}$ *m/z* = 1131.63, $[M+7H]^{7+}$ *m/z* = 1293.24, $[M+6H]^{6+}$ *m/z* = 1508.22, $[M+5H]^{5+}$ *m/z* = 1809.56.

**[Ala$^{183}$]-Cp183$^{1–183}$ (14).**

*Convergent Route*: Thioester **5** (10 mg, 0.78 µmol) and cysteine-peptide **13** (9.0 mg, 0.80 µmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 0.50 mL) for 20 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide **14** (1.7 mg, 9.0% yield).

*C-to-N Route*: Thioester **1** (0.97 mg, 0.13 µmol) and cysteine-peptide **16** (1.3 mg, 0.089 µmol) were reacted in ligation buffer (400 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 0.13 mL) for 2 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide **14** (0.61 mg, 32% yield). MS (ESI): calcd for

$C_{935}H_{1474}N_{278}O_{265}S_5$: 21010.06; observed: $[M+15H]^{15+}$ *m/z* = 1402.08, $[M+14H]^{14+}$ *m/z* = 1502.13, $[M+13H]^{13+}$ *m/z* = 1617.73, $[M+12H]^{12+}$ *m/z* = 1752.00, $[M+11H]^{11+}$ *m/z* = 1912.00.

**Cp183$^{61–183}$ (16).** MeNbz-peptide **2b** (1.2 mg, 0.17 μmol) and cysteine-peptide **13** (1.6 mg, 0.14 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 0.28 mL) for 3 h at 37 °C. Then, the same amount of deThz buffer (1 M methoxyamine hydrochloride, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 4.0; 0.28 mL) was added and the reaction mixture was incubated for 19 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide **16** (0.57 mg, 24% yield). MS (ESI): calcd for $C_{631}H_{1013}N_{201}O_{174}S_3$: 14295.46; observed: $[M+11H]^{11+}$ *m/z* = 1300.95, $[M+10H]^{10+}$ *m/z* = 1431.64, $[M+9H]^{9+}$ *m/z* = 1589.93.

**CD Analysis of Synthetic Cp183(C183A).** Lyophilized Cp183(C183A) peptide **14** was dissolved in denaturing buffer (6 M guanidine·HCl, 50 mM HEPES, 2 mM DTT, pH 7.5) at a concentration of ~0.2 mg/mL, and the mixture was incubated overnight at 4 °C. The denatured protein solution was dialyzed against 200 volumes of dialysis buffer (1.5 M guanidine·HCl, 50 mM HEPES, 2 mM DTT, pH 7.5) overnight at 4 °C. The protein solution was then further dialyzed against 200 volumes of various buffer conditions (1.0 M guanidine·HCl [or 0.5 M guanidine·HCl or 1.5 M NaCl], 50 mM HEPES, 2 mM DTT, pH 7.5) overnight at 4 °C. CD spectra of proteins were recorded using a JASCO J-720 circular dichroism spectrometer (JASCO, Tokyo, Japan) at 20 °C.

**Folding and Assembly of Capsids from Synthetic Cp183(C183A).** Lyophilized Cp183(C183A) peptide **9** was dissolved in denaturing buffer (6 M guanidine·HCl, 50 mM HEPES, 2 mM DTT, pH 7.5) at a concentration of ~0.2 mg/mL, and the mixture was incubated overnight at 4 °C. The denatured protein solution was dialyzed against 200 volumes of dialysis buffer (1.5 M guanidine·HCl, 50 mM HEPES, 2 mM DTT, pH 7.5) overnight at 4 °C. UV absorption spectra were obtained using a V-630 BIO spectrophotometer (JASCO, Tokyo, Japan). The concentration of Cp183(C183A) was calculated using the extinction coefficient of 58,900 $M^{-1}cm^{-1}$ (280 nm) per Cp183(C183A) dimer.[42] Then, 3 μL of 10 μM ssDNA (5'-ATGAATAACCAACGAAAAAAGGCGAGAAATACGCCTTTCAATATGCTGAA-3'; Thermo Fisher Scientific, Tokyo, Japan) was added to 7.5 μL of 3.1 μM Cp183(C183A). After incubating overnight at 4 °C, 12 μL $H_2O$ was added to the protein solution. The final capsid solution (1.0 μM Cp183(C183A), 13 μM ssDNA, 0.5 M

guanidine·HCl, 17 mM HEPES, 0.67 mM DTT, pH 7.5) was analyzed by transmission electron microscopy.

**Electron Microscopy.** Copper mesh grids coated with formvar and carbon (Veco grids, Nisshin EM, Tokyo, Japan) were glow-discharged and placed on drops of the specimen for 1 min, rinsed with distilled water, stained with a 2% uranyl acetate solution and examined with a transmission electron microscope (HT7700, Hitachi Ltd., Japan) at 80 kV.

References


1. Yuen, M.-F.; Chen, D.-S.; Dusheiko, G. M.; Janssen, H. L. A.; Lau, D. T. Y.; Locarnini, S. A.; Peters M. G.; Lai, C.-L. *Nat. Rev. Dis. Primers* **2018**, *4*, 18035.
2. WHO. Hepatitis B. Key facts. July 18, 2023. https://www.who.int/news-room/fact-sheets/detail/hepatitis-b (accessed November 2023).
3. Nassal, M. *Gut* **2015**, *64*, 1972–1984.
4. Dusheiko, G.; Agarwal, K.; Maini, M. K. *N. Engl. J. Med.* **2023**, *388*, 55–69.
5. Nassal, M. *J. Virol.* **1992**, *66*, 4107–4116.
6. Tan, Z.; Pionek, K.; Unchwaniwala, N.; Maguire, M. L.; Loeb, D. D.; Zlotnick, A. *J. Virol.* **2015**, *89*, 3275–3284.
7. King, R. W.; Ladner, S. K.; Miller, T. J.; Zaifert, K.; Perni, R. B.; Conway, S. C.; Otto, M. J. *Antimicrob. Agents Chemother.* **1998**, *42*, 3179–3186.
8. Deres, K.; Schröder, C. H.; Paessens, A.; Goldmann, S.; Hacker, H. J.; Weber, O.; Krämer, T.; Niewöhner, U.; Pleiss, U.; Stoltefuss, J.; Graef, E.; Koletzki, D.; Masantschek, R. N. A.; Reimann, A.; Jaeger, R.; Groß, R.; Beckermann, B.; Schlemmer, K.-H.; Haebich, D.; Rübsamen-Waigmann, H. *Science* **2003**, *299*, 893–896.
9. Campagna, M. R.; Liu, F.; Mao, R.; Mills, C.; Cai, D.; Guo, F.; Zhao, X.; Ye, H.; Cuconati, A.; Guo, H.; Chang, J.; Xu, X.; Block, T. M.; Guo, J.-T. *J. Virol.* **2013**, *87*, 6931–6942.
10. Kang, J.-A.; Kim, S.; Park, M.; Park, H.-J.; Kim, J.-H.; Park, S.; Hwang, J.-R.; Kim, Y.-C.; Kim, Y.-J.; Cho, Y.; Jin, M. S.; Park, S.-G. *Nat. Commun.* **2019**, *10*, 2184.
11. Stray, S. J.; Bourne, C. R.; Punna, S.; Lewis, W. G.; Finn, M. G.; Zlotnick, A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 8138–8143.
12. Feld, J. J.; Colledge, D.; Sozzi, V.; Edwards, R.; Littlejohn, M.; Locarnini, S. A. *Antiviral Res.* **2007**, *76*, 168–177.
13. Viswanathan, U.; Mani, N.; Hu, Z.; Ban, H.; Du, Y.; Hu, J.; Chang, J.; Guo, J.-T. *Antiviral Res.* **2020**, *182*, 104917.
14. Harrison, K.; Mackay, A. S.; Kambanis, L.; Maxwell, J. W. C.; Payne, R. J. *Nat. Rev. Chem.* **2023**, *7*, 383–404.
15. Bourne, C. R.; Finn, M. G.; Zlotnick, A. *J. Virol.* **2006**, *80*, 11055–11061.
16. Zlotnick, A.; Venkatakrishnan, B.; Tan, Z.; Lewellyn, E.; Turner, W.; Francis, S. *Antiviral Res.* **2015**, *121*, 82–93.
17. Gallina, A.; Bonelli, F.; Zentilin, L.; Rindi, G.; Muttini, M.; Milanesi, G. *J. Virol.* **1989**, *63*, 4645–4652.

18. Wingfield, P. T.; Stahl, S. J.; Williams, R. W.; Steven, A. C. *Biochemistry* **1995**, *34*, 4919–4932.
19. Ceres, P.; Zlotnick, A. *Biochemistry* **2002**, *41*, 11525–11531.
20. Tsuda, S.; Mochizuki, M.; Ishiba, H.; Yoshizawa-Kumagaye, K.; Nishio, H.; Oishi, S.; Yoshiya, T. *Angew. Chem., Int. Ed.* **2018**, *57*, 2105–2109.
21. Porterfield, J. Z.; Dhason, M. S.; Loeb, D. D.; Nassal, M.; Stray, S. J.; Zlotnick, A. *J. Virol.* **2010**, *84*, 7174–7184.
22. Nassal, M.; Rieger, A.; Steinau, O. *J. Mol. Biol.* **1992**, *225*, 1013–1025.
23. Birnbaum, F.; Nassal, M. *J. Virol.* **1990**, *64*, 3319–3330.
24. Fujiwara, Y.; Akaji, K.; Kiso, Y. *Chem. Pharm. Bull.* **1994**, *42*, 724–726.
25. Lukszo, J.; Patterson, D.; Albericio, F.; Kates, S. A. *Lett. Pept. Sci.* **1996**, *3*, 157–166.
26. Wang, J.-X.; Fang, G.-M.; He, Y.; Qu, D.-L.; Yu, M.; Hong, Z.-Y.; Liu, L. *Angew. Chem., Int. Ed.* **2015**, *54*, 2194–2198.
27. Tsuda, S.; Masuda, S.; Yoshiya, T. *Org. Biomol. Chem.* **2019**, *17*, 1202–1205.
28. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.
29. Tan, Y.; Wu, H.; Wei, T.; Li, X. *J. Am. Chem. Soc.* **2020**, *142*, 20288–20298.
30. Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248–9252.
31. Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O. *Chem. Rev.* **2019**, *119*, 7328–7443.
32. Fuchs, O.; Trunschke, S.; Hanebrink, H.; Reimann, M.; Seitz, O. *Angew. Chem., Int. Ed.* **2021**, *60*, 19483–19490.
33. Hojo, H.; Murasawa, Y.; Katayama, H.; Ohira, T.; Nakahara, Y.; Nakahara, Y. *Org. Biomol. Chem.* **2008**, *6*, 1808–1813.
34. Blanco-Canosa, J. B.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2008**, *47*, 6851–6855.
35. Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc.* **2015**, *137*, 7197–7209.
36. Bang, D.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534–2538.
37. Crowther, R. A.; Kiselev, N. A.; Böttcher, B.; Berriman, J. A.; Borisova, G. P.; Ose, V.; Pumpens, P. *Cell* **1994**, *77*, 943–950.
38. Bin Mohamed Suffian, I. F.; Garcia-Maya, M.; Brown, P.; Bui, T.; Nishimura, Y.; Bin Mohammad Johari Palermo, A. R.; Ogino, C.; Kondo, A.; Al-Jamal, K. T. *Sci. Rep.* **2017**, *7*, 43160.
39. Machida, S.; Ogawa, S.; Xiaohua, S.; Takaha, T.; Fujii, K.; Hayashi, K. *FEBS Lett.* **2000**, *486*, 131–135.
40. T. Noguchi, S. Oishi, K. Honda, Y. Kondoh, T. Saito, H. Ohno, H. Osada and N. Fujii,

*Chem. Commun.* **2016**, *52*, 7653–7656.
41. Aoki, K.; Higashi, K.; Oda, S.; Manabe, A.; Maeda, K.; Morise, J.; Oka, S.; Inuki, S.; Ohno, H.; Oishi, S.; Nonaka, M. Submitted.
42. Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. *Protein Sci.* **1995**, *4*, 2411–2423.

Chapter 2. Synthetic Studies of Therapeutic Targets for Mirror-Image Screening

Section 2. Chemical Synthesis of Interleukin-6 and Its Application to Screening of a Mirror-Image VHH Library

*Summary: Interleukin-6 (IL-6), a multifunctional cytokine, is an attractive therapeutic target for immunological and inflammatory diseases. The author investigated the chemical synthesis of IL-6 and its enantiomer (D-IL-6) using a sequential N-to-C native chemical ligation strategy. Solubilizing Trt-$K_{10}$ tags, which were used to facilitate the synthesis of full-length IL-6 protein from six peptide segments, improved the intermediate solubility and served as protecting groups during the metal-free desulfurization. Synthetic L-IL-6 and recombinant L-IL-6 exhibited identical structural and binding properties. The symmetrical binding property of D-IL-6 was also demonstrated by functional analysis using IL-6-binding peptides. The resulting functional D-IL-6 was employed to screen a phage-displayed antibody fragment library, leading to the identification of several D-IL-6-binding sequences with high homology.*

Interleukin-6 (IL-6, **1**) is a pleiotropic cytokine that functions in various biological systems, including tissue regeneration, immunity, and metabolism.[1–4] The IL-6 signaling cascade is initiated by the binding of IL-6 to its cognate receptor IL-6R and a second transmembrane protein, glycoprotein (gp)130. This leads to activation of the Janus kinase/signal transducer and activator of transcription 3 (JAK-STAT3) pathway, the mitogen-activated protein kinase (MAPK) cascade, and the phosphatidylinositide-3-kinase (PI3K) cascade.[5,6] Because excessive IL-6 production is associated with several immunological and inflammatory disorders, including rheumatoid arthritis, Castleman disease, and Crohn's disease, the IL-6 pathway is considered to be a promising therapeutic target.[7]

To date, several therapeutic antibodies that block IL-6 signaling have been approved for treatment. A humanized anti-human IL-6R monoclonal antibody, tocilizumab, has achieved significant improvements in several diseases, including rheumatoid arthritis.[8] An anti-human IL-6 monoclonal antibody, siltuximab,[9] was approved for the treatment of Castleman disease, and several antibodies targeting IL-6 are now in clinical trials.[10] Small protein scaffolds that bind to IL-6 with sub-nanomolar affinity have also been developed, including 18-kDa avimers,[11] 6-kDa affibodies,[12] and variable domain of the heavy chain antibody (VHH) derived from *Camelidae*.[13] For example, bispecific VHH,

which binds to both IL-6 and tumor necrosis factor (TNF), is indicated for the high-potency treatment of rheumatoid arthritis in both *in vitro* and *in vivo* models.[13]

The long-term administration of protein-based therapeutics targeting IL-6 signaling can induce undesirable immune responses that lead to adverse effects and/or neutralization of drug activity.[14] The gradual decrease in therapeutic effects has been a major impediment to durable, long-term clinical response in the treatment of rheumatic diseases.[15,16] For example, it was reported that the administration of sarilumab, an approved anti-IL-6R human monoclonal antibody, caused anti-drug antibody (ADA) generation in approximately 10% of patients.[17] In the phase IIb trial of ALX-0061, a humanized VHH dimer targeting IL-6R, ADA generation was observed in over 30% of the ALX-0061-treated patients.[18] Thus, successful development of novel therapeutic agents targeting IL-6 cascades with long-lasting clinical effects will require overcoming these drawbacks of potential immunogenic response.

As described in Preface, mirror-image proteins (D-proteins) applied as biomaterials effectively circumvent the potential risk of immunogenicity and rapid degradation.[19] In this Section, the author describes the chemical synthesis of a mirror-image protein of IL-6, which is applicable to mirror-image screening (Figure 1). Comparative analysis of both enantiomers of IL-6 was also performed to characterize their structures and functions.

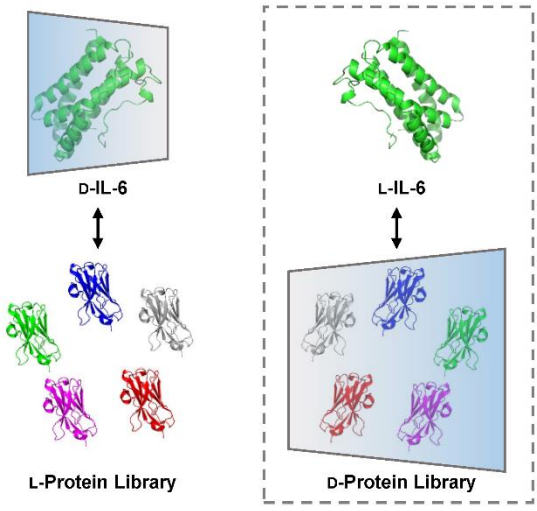


**Figure 1.** Mirror-image screening process. Screening of an L-protein library using D-IL-6 corresponds to that of a D-protein library using L-IL-6 in a mirror. The protein structures shown here represent library proteins and were obtained from Protein Data Bank (PDBID: 1ALU [IL-6]; 5TP3).

**Design of the Synthetic Strategy and Initial Synthetic Attempt**. IL-6 (**1**) is a 183-residue protein with two disulfide bonds at cysteine (Cys)$^{43}$-Cys$^{49}$ and Cys$^{72}$-Cys$^{82}$ (Figure 2A).[20] IL-6 has an N-glycan at asparagine (Asn)$^{44}$ and multiple O-glycans, but glycosylation of human IL-6 does not affect its binding to IL-6R and gp130.[21] The less significant contribution of the N-glycan in IL-6 was also revealed by structure-activity relationship studies of a series of glycosylated forms, which were prepared by semi-synthesis from recombinant proteins and synthetic glycopeptides (Figure 2B).[22,23] Initially, the author designed a synthetic route for non-glycosylated form of L-IL-6 using a sequential native chemical ligation (NCL) strategy.[24] Previous reports of semi-synthesis of glycosylated IL-6[22,23] cited concerns about the low solubility of peptide segment(s) and the labile aspartic acid (Asp)$^{139}$-proline (Pro)$^{140}$ bond under acidic conditions. For synthetic studies of IL-2 having a hydrophobic sequence at the C-terminus, several temporary modification, including isopeptide bonds and solubilizing tags, were also employed.[25–27] To avoid the potential low solubility of the C-terminal segment of IL-6, the author used the solubilizing Trt-$K_{10}$ tag, which comprises a hydrophilic oligo-lysine (Lys) sequence and a trityl anchor.[28] Trt-$K_{10}$ can be selectively attached and detached on the side chain thiol of a Cys residue to serve as a temporary protecting group during metal-free desulfurization (MFD), which converts the precursor Cys residue(s) into alanine

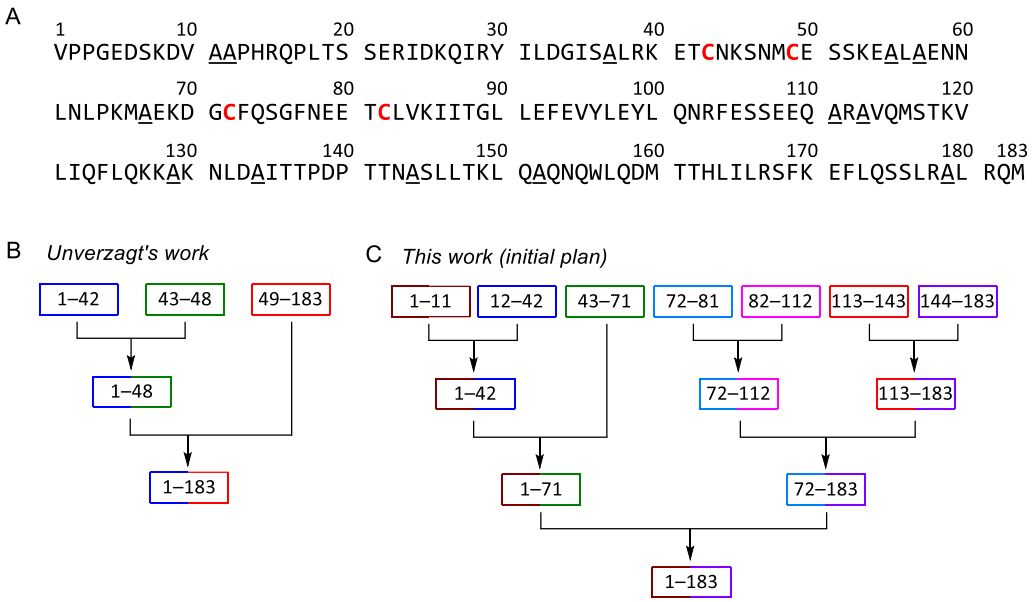


**Figure 2.** Sequence and synthetic strategy for IL-6. (A) Sequence of IL-6. Red-colored letters: Cys residues; underlined letters: Ala residues. (B) Strategy for semi-synthesis of IL-6 glycoforms by Unverzagt *et al.*[22,23] N-terminal (blue) and C-terminal (red) segments were prepared by recombinant expression, and the middle segment (green) was prepared by chemical synthesis. A series of N-glycans was introduced at Asn$^{44}$. (C) Initial plan for chemical synthesis of IL-6$^{His}$ in this study.

(Ala).[29,30] The author also attached a poly-histidine (His) tag at the N-terminus of the sequence via hydrophilic polyethylene glycol (PEG) linker to immobilize IL-6 during the screening step. In the initial attempt, the author planned to synthesize L-IL-$6^{His}$ (L-**1a**) via a convergent ligation strategy from seven peptide segments, which were split at $Ala^{12}$, $Cys^{43}$, $Cys^{72}$, $Cys^{82}$, $Ala^{113}$, and $Ala^{144}$ of IL-6 (Figure 2C).

Several excellent latent thioesters have been developed for NCL, including the first-generation Dawson linker (Dbz),[31] hydrazide,[32] and SEAlide system.[33] These latent thioesters can prevent the undesirable oligomerization and intramolecular cyclization of peptides in the N-to-C direction during NCL because of the low reactivity before activation.[34] Of these options, the author selected the Dbz linker because the incorporation of a solubilization sequence into the latent thioester would be advantageous.[35] To improve the solubility of peptide segments, four arginine (Arg) residues were attached downstream of the Dbz linker.[36] Fmoc-based solid-phase peptide synthesis (Fmoc-SPPS) was used to synthesize all of the following peptide segments: IL-$6^{1–11}$ with poly-histidine (His) tag (L-**2a**), IL-$6^{12–42}$ (L-**3**), IL-$6^{43–71}$ (L-**4**), IL-$6^{72–81}$ (L-**5**), IL-$6^{82–112}$ (L-**6**), IL-$6^{113–143}$ (L-**7**), and IL-$6^{144–183}$ (L-**8**) (Figures 3). Dbz-peptides L-**2a**, L-**3**, and L-**6** were smoothly synthesized by standard Fmoc-SPPS. Unfavorable aspartimide formation during the elongation of peptide L-**4** containing $Asp^{70}$-glycine (Gly)$^{71}$ sequences at the C-terminus was observed in liquid chromatography-mass spectrometry (LC-MS) analysis.[37] To minimize the potential for aspartimide formation, Fmoc deprotection was performed here under mild conditions [5% piperazine-0.1 M Oxyma Pure in *N*-methyl-2-pyrrolidone (NMP)/EtOH] (Figure 4).[38] Peptide L-**5** was also synthesized by the standard SPPS approach, which contains acetamidomethyl (Acm)-protected Cys for temporary N-terminal protection and an *o*-amino(methyl)aniline (MeDbz) linker[39] for C-terminal activation. Peptide thioester L-**7** was synthesized via a side-chain-anchoring strategy[40] on Rink amide resin to avoid the undesirable intramolecular cyclization of C-terminal Asn.[32,34] C-terminal peptide L-**8**

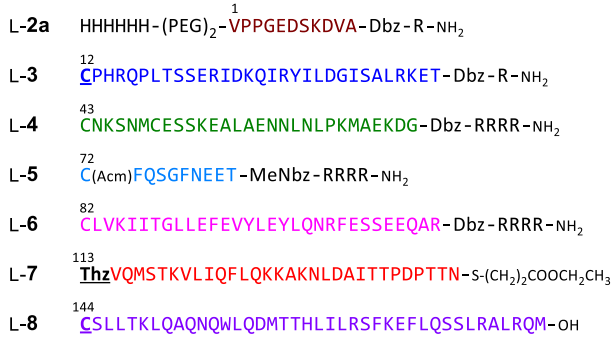


**Figure 3.** Amino acid sequences of the initially planned peptide segments for IL-$6^{His}$ .

was synthesized by a standard protocol for microwave-assisted SPPS[38] because it does not contain sequences that are sensitive to harsh conditions.

With the originally planned peptide segments (L-**2a**–**8**) in hand, the author investigated the sequential ligations. Dbz-peptide L-**2a** was converted to thioester L-**9a** by $NaNO_2$-mediated activation followed by the addition of 4-mercaptophenylacetic acid (MPAA)

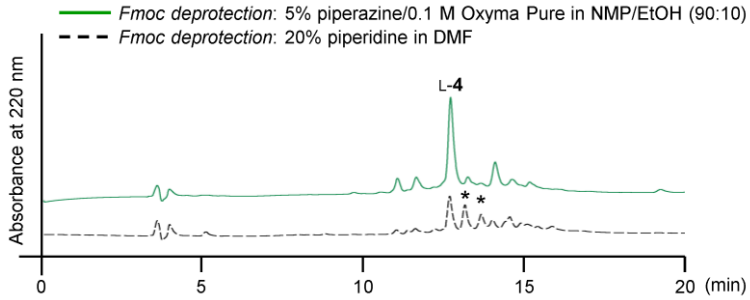


**Figure 4.** Optimization of Fmoc-deprotection conditions for the synthesis of peptide L-**4**. Overlay of the chromatograms of reaction mixtures of Fmoc-deprotection. * indicates the aspartimide-related byproduct. *HPLC conditions*. Cosmosil 5C18-AR300 column (Nacalai Tesque, 4.6 × 250 mm), linear gradient of 15–35% $CH_3CN$ containing 0.05% TFA at a flow rate of 1 mL/min over 20 min.

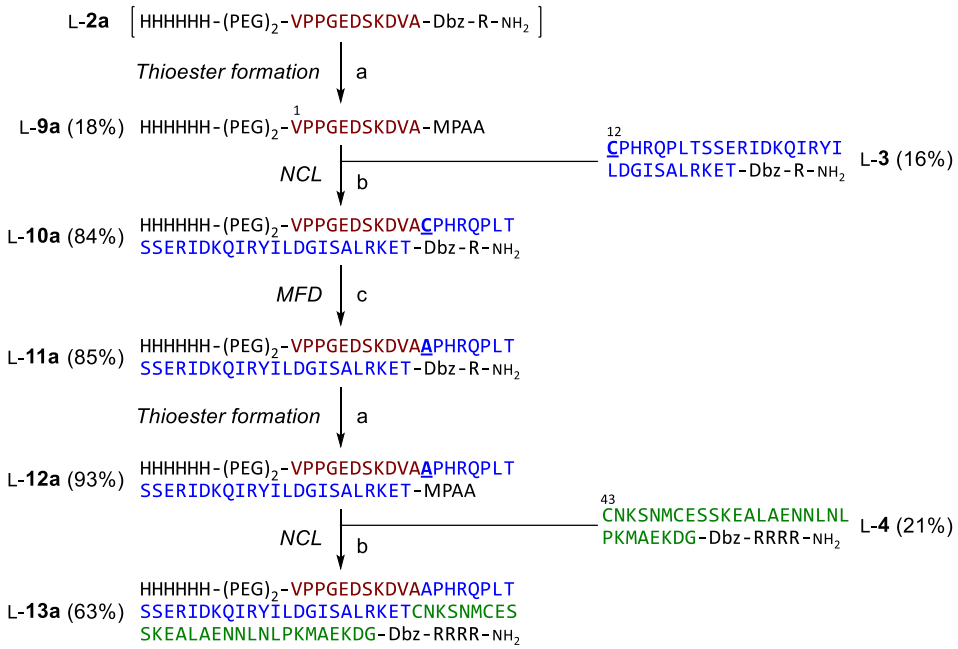


**Figure 5.** Synthesis of N-terminal IL-$6^{1–71}$ (L-**13a**). *Reagents and conditions*: (a) $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, then MPAA and TCEP; (b) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (c) VA-044, MESNa, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 6.5).

(Figure 5).[41] Peptide thioester L-**9a** and peptide L-**3** were ligated in the presence of MPAA to provide peptide L-**10a**. The temporarily mutated Cys$^{12}$ in L-**10a** was converted to Ala$^{12}$ by radical-mediated MFD with VA-044[29] to afford the Ala$^{12}$ congener L-**11a** (IL-6$^{1–42}$). $NaNO_2$-mediated activation of the C-terminal Dbz linker in L-**11a** followed by treatment with MPAA provided thioester L-**12a**. The subsequent NCL between thioester L-**12a** and peptide L-**4** in the presence of MPAA gave the N-terminal IL-6$^{1–71}$ (L-**13a**) in 63% yield.

For the synthesis of the middle segment L-**15**, MeNbz peptide L-**5** and peptide L-**6** were subjected to 1,2,4-triazole-mediated NCL conditions (Figure 6). Subsequently, free Cys$^{82}$ was protected by Trt(OH)-$K_{10}$ to afford the desired IL-6$^{72–112}$ (L-**14**) with a hydrophilic Trt-$K_{10}$ tag.[28] $NaNO_2$-mediated activation of L-**14** followed by the addition of MPAA provided the MPAA thioester L-**15**. For the synthesis of the C-terminal segment L-**18**, alkyl thioester L-**7** was converted to the MPAA thioester L-**16** by treatment with MPAA in slightly acidic buffer. Notably, the appropriate pH adjustment was essential for this reaction to minimize the undesired hydrolysis of the thioester. NCL between thioester

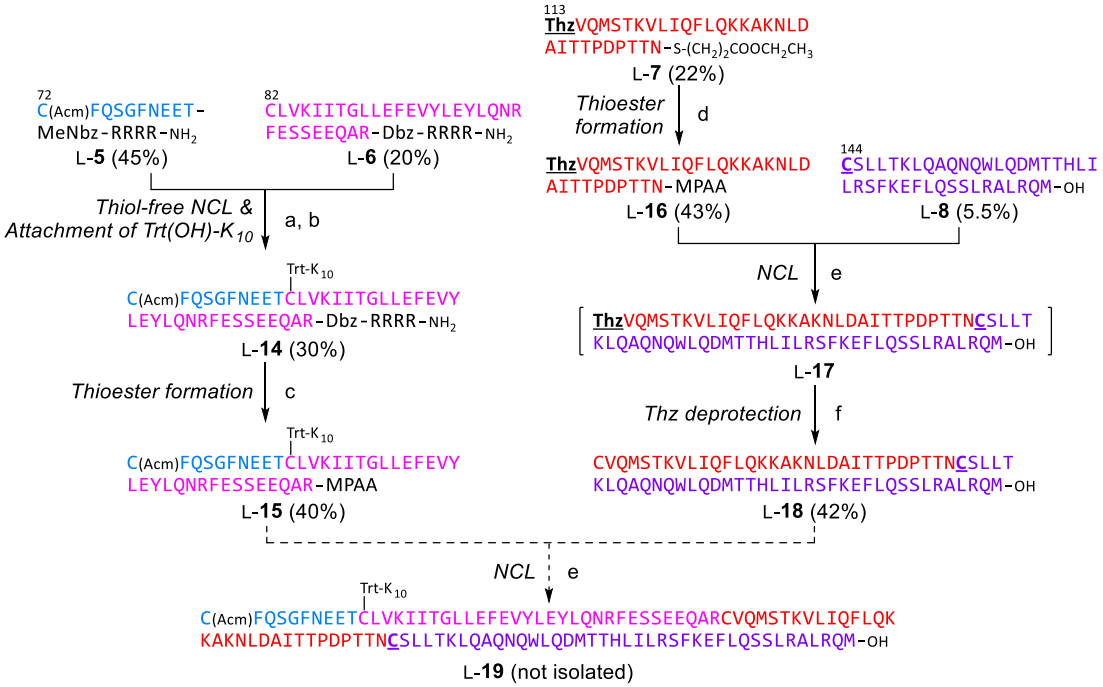


**Figure 6.** Initial attempts to synthesize IL-6$^{72–183}$ (L-**19**). Intermediate peptide L-**19** was not isolated because only a trace amount of L-**19** was eluted from the HPLC column under the standard conditions. *Reagents and conditions*: (a) 1,2,4,-triazole, TCEP, 6 M guanidine·HCl, and 100 mM phosphate buffer (pH 7.1); (b) Trt(OH)-$K_{10}$ and TFA; (c) $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, then MPAA and TCEP; (d) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 6.5); (e) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (f) methoxyamine.

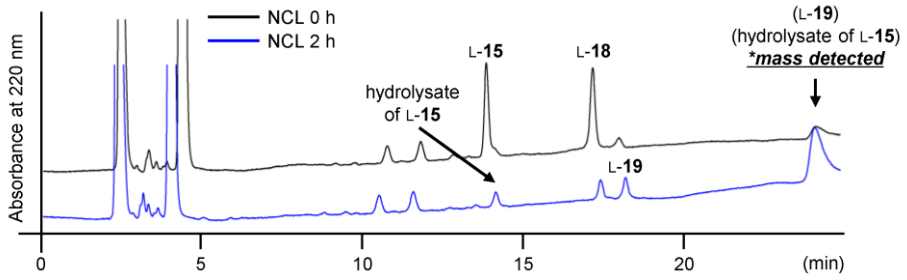


**Figure 7.** HPLC trace of the initial attempt at NCL to synthesize C-terminal segment L-**19** in the convergent route. *HPLC conditions*. Cosmosil 5C4-AR300 column (Nacalai Tesque, 4.6 × 150 mm), linear gradient of 35–55% $CH_3CN$ containing 0.05% TFA at a flow rate of 1 mL/min over 20 min at 40 °C (0–20 min: 35–55%; 20–22 min: 95%).

L-**16** and peptide L-**8** provided thiazolidine (Thz)-protected L-**17**. Subsequent deprotection of the Thz moiety by the addition of methoxyamine[42] afforded the desired C-terminal segment L-**18** (IL-$6^{113–183}$). Next, middle segment L-**15** and C-terminal segment L-**18** were subjected to standard MPAA-mediated NCL conditions. However, after the consumption of starting materials, only a trace amount of the desired ligation product L-**19** (IL-$6^{72–183}$) was detected by high-performance liquid chromatography (HPLC) analysis. Most of the ligation product was not eluted from the analytical HPLC column under the standard conditions probably because of its low solubility (Figure 7).

**Preparation of Synthetic L-IL-6 via a Modified Protocol**. Next, the author investigated the synthesis of IL-$6^{His}$ (L-**1a**) by an N-to-C NCL strategy to avoid the use of IL-$6^{72–183}$ (L-**19**), which had low solubility (Figure 8). The author expected that the N-terminal His tag and PEG linker would enhance the solubility of intermediate peptide segments. Because the N-terminal and C-terminal peptide segments L-**13a** and L-**18** could be used for the N-to-C NCL strategy, the author only designed and synthesized one new middle segment (L-**20**), which contains a Dbz linker at the C-terminus for late-stage activation. The $NaNO_2$-mediated thioester formation of L-**13a** followed by NCL reaction with L-**20** provided IL-$6^{1–112}$ (L-**21a**). Next, four Cys residues in L-**21a** ($Cys^{43}$, $Cys^{49}$, $Cys^{72}$, and $Cys^{82}$) were protected by treatment with an excess amount of Trt(OH)-$K_{10}$ reagent[28] to provide Trt-$K_{10}$-modified N-terminal segment L-**22a** in 71% yield. The thioester formation of L-**22a** followed by NCL reaction with L-**18** afforded the Trt-$K_{10}$-modified full-length peptide L-**23a** in sufficient yield. VA-044-mediated desulfurization of L-**23a** proceeded smoothly without detectable side reactions to provide the desulfurized peptide L-**24a** in 54% yield. For the synthesis of L-**23a** and L-**24a**, the author carefully adjusted the pH of the reaction buffer to allow the reaction to proceed efficiently, given

the formation of trifluoroacetic acid (TFA) salts with basic amino acids in the sequences containing multiple Trt-$K_{10}$ protecting groups. Finally, four Trt-$K_{10}$ protecting groups were cleaved by treatment with TFA/triisopropylsilane (TIS)[28] to provide the reduced form of the desired full-length His-tagged IL-$6^{1–183}$ (L-**25a**).

For the preparation of IL-6 and other interleukins, appropriate folding to the oxidized form with disulfide bonds is critical to reproduce biological function.[20,24,43] Therefore,

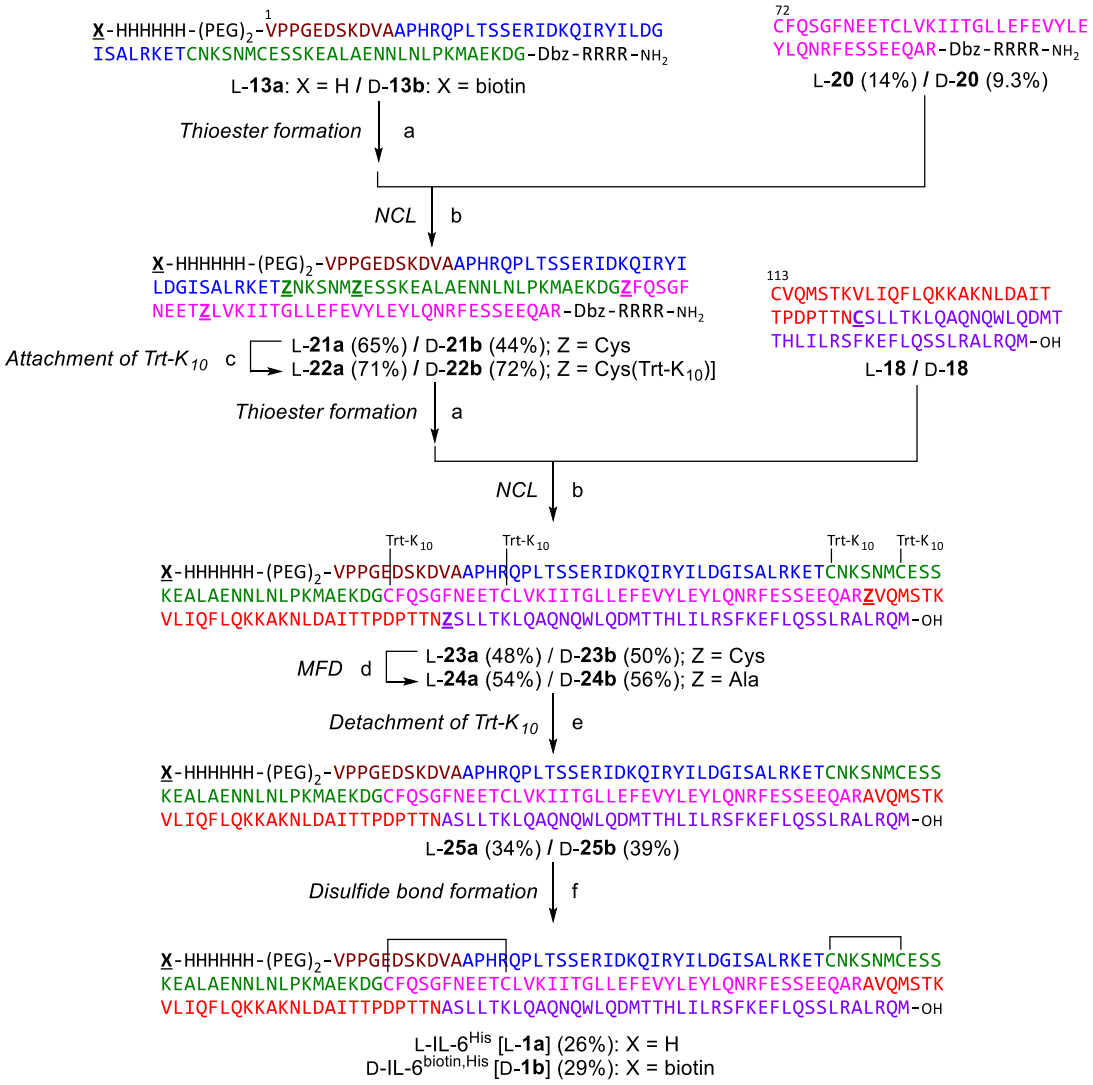


**Figure 8.** Synthesis of L-IL-$6^{His}$ (L-**1a**) and D-IL-$6^{biotin,His}$ (D-**1b**) via N-to-C NCL. *Reagents and conditions*: (a) $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, then MPAA and TCEP; (b) MPAA, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 7.0); (c) Trt(OH)-$K_{10}$ (for L-**22a**) or Trt(OH)-$k_{10}$ (for D-**22b**), 6 M guanidine·HCl, and 200 mM phosphate buffer/TFA (1:1); (d) VA-044, MESNa, TCEP, 6 M guanidine·HCl, and 200 mM phosphate buffer (pH 6.5); (e) TFA/TIS (95:5); (f) 6 M guanidine, 10 mM Tris (pH 8.5), then 10-fold dilution into 10 mM Tris (pH 8.5).

using the folding conditions of recombinant IL-6,[44] the resulting synthetic L-IL-6$^{His}$ (reduced form) (L-**25a**) was subjected to oxidation and refolding to obtain the functional protein. After lyophilized L-IL-6$^{His}$ (reduced form) (L-**25a**) was denatured in 6 M guanidine, the mixture was diluted 10-fold in Tris buffer (pH 8.0). A change in the retention time of reverse-phase HPLC was observed after incubation for 21 h (Figure 9). Size-exclusion chromatography (SEC) analysis revealed that the retention time of the main peak of folded synthetic product was identical to that of recombinant IL-6 (Figure 10). The author then purified folded L-IL-6$^{His}$ by SEC for further structural and binding analyses.

To confirm the formation of disulfide bonds at appropriate positions, the author used

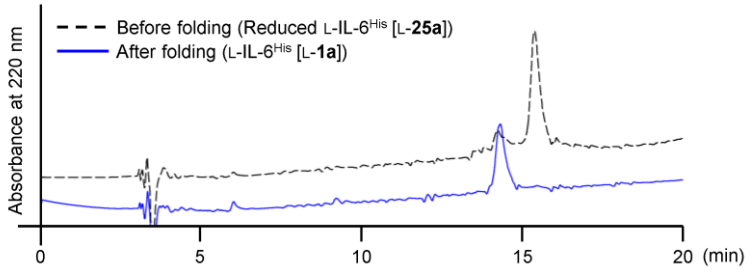


**Figure 9.** Conversion of reduced L-IL-6$^{His}$ (L-**25a**) into folded L-IL-6$^{His}$ (L-**1a**). Overlay of the HPLC chromatograms before (reduced L-IL-6$^{His}$) and after (L-IL-6$^{His}$) folding. *HPLC conditions*: Cosmosil 5C18-AR300 column (Nacalai Tesque, 4.6 × 250 mm), linear gradient of 40–60% $CH_3CN$ containing 0.05% TFA at a flow rate of 1 mL/min over 20 min.

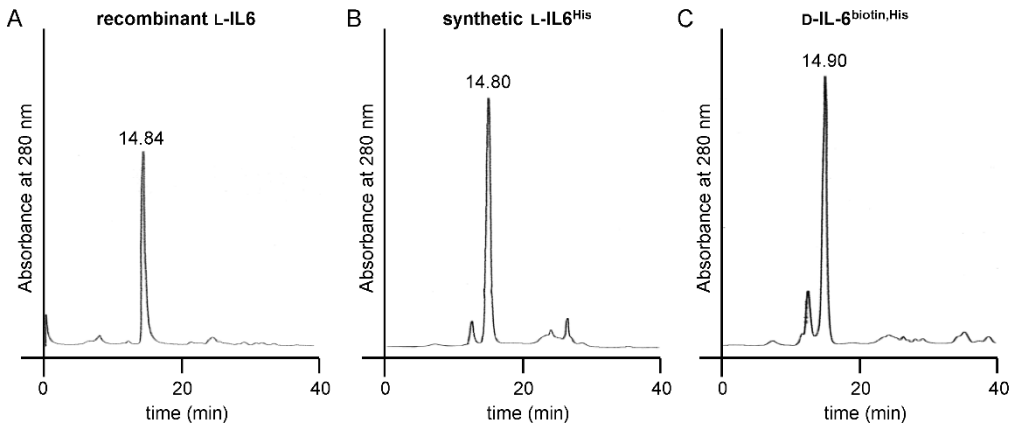


**Figure 10.** Purification of folded synthetic L-IL-6$^{His}$ and D-IL-6$^{biotin,His}$ by SEC. The main peak was purified to analyze CD spectra and binding activities. (A) Chromatogram of recombinant L-IL-6. (B) Chromatogram of synthetic L-IL-6$^{His}$ after folding reaction. (C) Chromatogram of D-IL-6$^{biotin,His}$ after folding reaction. *SEC conditions*: Superdex™ 75 Increase 10/300 GL column (Cytiva), PBS (pH 7.4) at a flow rate 0.8 mL/min.

mass spectrometry to analyze peptide fragments derived from protease-mediated digestion.[44,45] Digestion of folded L-IL-6$^{His}$ by Lys-C protease under non-reducing conditions generated peptide fragments, which were cleaved at the C-termini of Lys residues (Figure 11A). The observed *m/z* values derived from the digested fragments were well matched with the theoretical *m/z* values containing disulfide bonds at Cys$^{43}$-Cys$^{49}$ and Cys$^{72}$-Cys$^{82}$. Additionally, peptide fragments derived from the product(s) with scrambled disulfide bonds were not detected. In contrast, dithiothreitol (DTT)-mediated reduction of cystines in folded L-IL-6$^{His}$, iodoacetamide-mediated *S*-alkylation, and

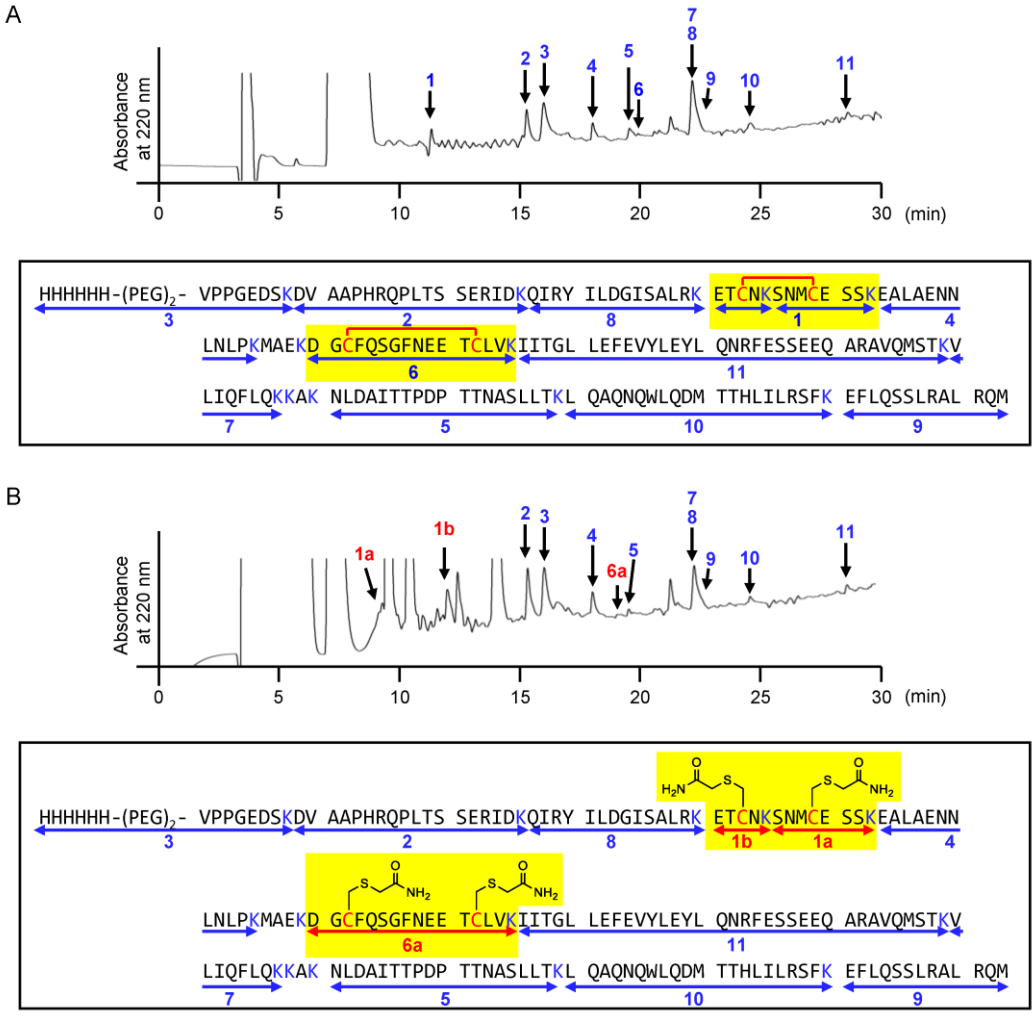


**Figure 11.** Analysis of disulfide bond formation in folded synthetic L-IL-6$^{His}$ (L-**1a**). After iodoacetamide-mediated alkylation of L-**1a** under non-reducing (A) or reducing (B) conditions, the mixture was subjected to Lys-C protease-mediated digestion. Peptide fragments were identified by electrospray ionization (ESI)-MS analysis. *HPLC conditions*: Cosmosil 5C18-AR300 column (Nacalai Tesque, 4.6 × 250 mm), linear gradient of 0–60% $CH_3CN$ containing 0.05% TFA at a flow rate of 1 mL/min over 30 min.

protease digestion gave an alternative set of peptide fragments, for which the observed *m/z* values were identical to the theoretical *m/z* values of *S*-alkylated peptides (Figure 11B). These results indicated that folded L-IL-6$^{His}$ (L-**1a**), which was prepared by folding processes under oxidative conditions, contained the appropriate disulfide bonds at Cys$^{43}$-Cys$^{49}$ and Cys$^{72}$-Cys$^{82}$. Next, the author compared the circular dichroism (CD) spectra of synthetic L-IL-6$^{His}$ and recombinant L-IL-6 to confirm the structures of the folded proteins. The CD spectrum of synthetic L-IL-6$^{His}$ was consistent with that of recombinant L-IL-6,[22,23] with two negative bands at 222 and 208 nm derived from the typical α-helix structure (Figure 12). To validate the binding activity of synthetic L-IL-6$^{His}$ toward IL-6R, the author performed surface plasmon resonance (SPR) analysis of IL-6R immobilized on a CM5 sensor chip. The affinity of synthetic L-IL-6$^{His}$ ($K_D$ = 20.0 ± 2.8 nM) toward IL-6R was equipotent to that of recombinant L-IL-6 ($K_D$ = 24.0 ± 0.9 nM) (Figures 13A and 13B). These results demonstrated that the structural and binding properties of synthetic L-IL-6$^{His}$ and recombinant L-IL-6 were identical.

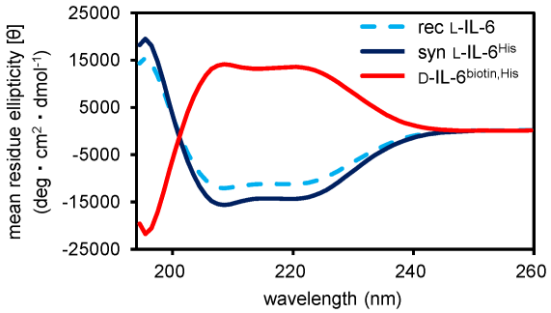


**Figure 12.** CD spectra of recombinant L-IL-6, synthetic L-IL-6$^{His}$ and D-IL-6$^{biotin,His}$.

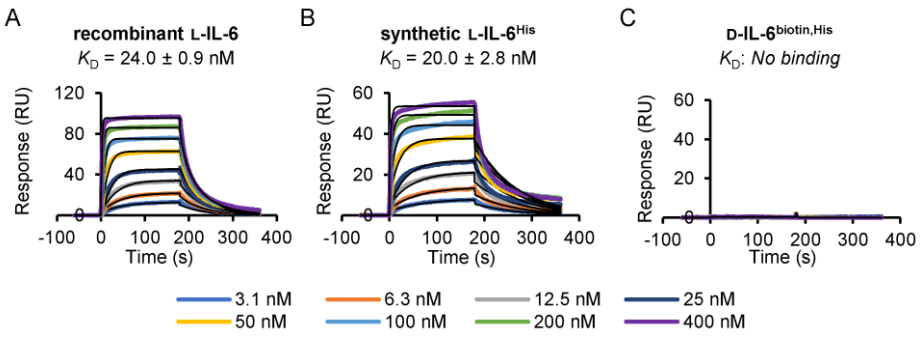


**Figure 13.** Representative SPR data analysis of recombinant L-IL-6 (A), synthetic L-IL-6$^{His}$ (B), and D-IL-6$^{biotin,His}$ (C) binding to the human IL-6R. Binding affinities were determined by fitting a 1:1 binding model from triplicate assays.

**Preparation and Enantiomeric Properties of D-IL-6.** For the synthesis of D-IL-6, biotin and D-His tags were incorporated at the N-terminus to enable interactions with streptavidin and $Ni^{2+}$, respectively, for immobilization in mirror-image screening (Figure 8). The Trt-$k_{10}$ tag, which comprises a D-Lys sequence, was also employed to identically reproduce the properties of the peptide segment and synthetic intermediates with those of materials used to synthesize L-IL-$6^{His}$. Although an additional biotin label was appended on the N-terminus of D-IL-6, each step of the synthetic process proceeded with comparable yields to those for L-IL-$6^{His}$, including NCL, attachment/detachment of solubilizing tags, and folding processes to produce folded D-IL-$6^{biotin,His}$ (D-**1b**). Folded D-IL-$6^{biotin,His}$ was eluted by SEC with a similar retention time to L-IL-$6^{His}$ (Figure 10). Symmetrical spectra of L-IL-$6^{His}$ and D-IL-$6^{biotin,His}$ were observed, suggesting that D-IL-$6^{biotin,His}$ has the mirror-image structure of native L-IL-6 (Figure 12).

Next, the author assessed the biological activity of the folded D-IL-$6^{biotin,His}$, which did not bind to IL-6R as expected (Figure 13C). Although, in principle, D-IL-$6^{biotin,His}$ should bind to D-IL-6R, the preparation of 80-kDa D-IL-6R was challenging. Thus, the

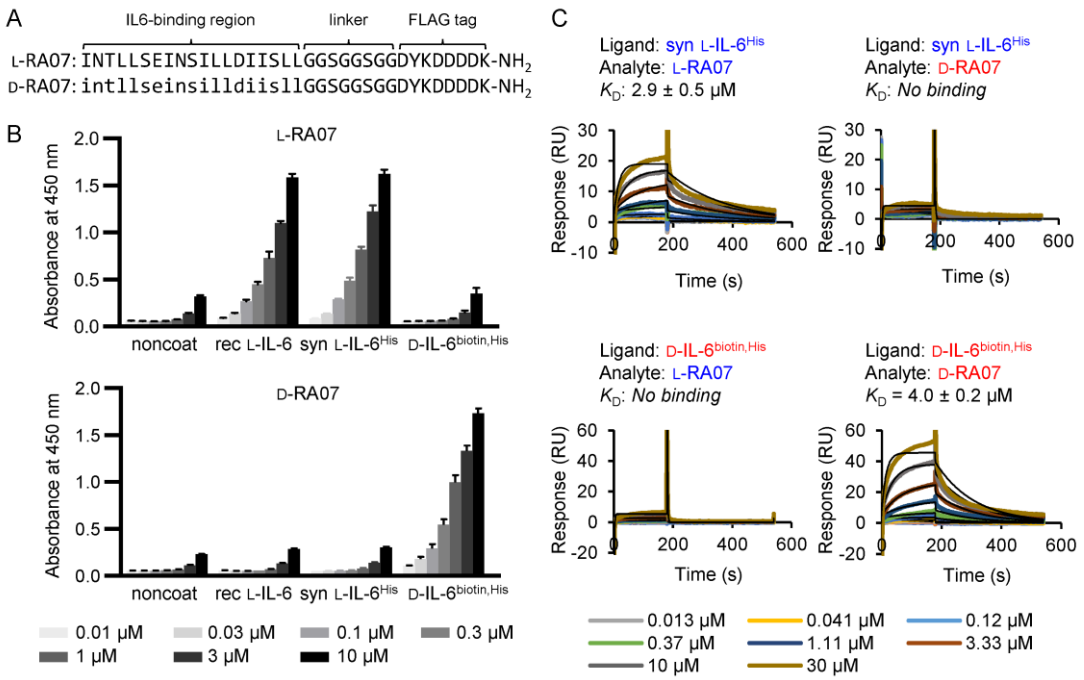


**Figure 14.** Functional analysis of synthetic L-IL-$6^{His}$ and D-IL-$6^{biotin,His}$. (A) Peptide sequences of L-RA07 and D-RA07. (B) ELISA of binding of recombinant L-IL-6, synthetic L-IL-$6^{His}$, and D-IL-$6^{biotin,His}$ towards L-RA07 (above) and D-RA07 (below). Mean absorbance values ± standard deviation are shown. (C) Representative SPR data for L-RA07 (left) and D-RA07 (right) binding to synthetic L-IL-$6^{His}$ (above) and D-IL-$6^{biotin,His}$ (below). Binding affinities were determined by fitting a 1:1 binding model from triplicate assays.

author focused on using an IL-6-binding peptide, RA07,[46] to evaluate the binding activity of both enantiomers of IL-6. RA07 is a 19-residue peptide that was identified by *in vitro* selection using mRNA display technology.[46] The author used Fmoc-SPPS to synthesize L-RA07 and D-RA07, which were labeled with a FLAG-tag attached to the RA07 C-terminus via a Gly-serine (Ser) linker to enable detection (Figure 14A). An enzyme-linked immunosorbent assay (ELISA) revealed that L-RA07 bound sufficiently to recombinant L-IL-6 and synthetic L-IL-6$^{His}$ but not to D-IL-6$^{biotin,His}$ (Figure 14B). In contrast, D-RA07 did not bind to recombinant L-IL-6 or synthetic L-IL-6$^{His}$, but did bind to D-IL-6$^{biotin,His}$, indicating that the binding activities of D-IL-6$^{biotin,His}$ and L-IL-6 are symmetrical. The binding affinities for RA07 of L-IL-6$^{His}$ and D-IL-6$^{biotin,His}$ immobilized on a CM5 sensor chip were determined by SPR analysis. The affinity of L-IL-6$^{His}$ with L-RA07 ($K_D$ = 2.9 ± 0.5 μM) was similar to that of D-IL-6$^{biotin,His}$ with D-RA07 ($K_D$ = 4.0 ± 0.2 μM), whereas no binding between L-IL-6$^{His}$ and D-RA07 or D-IL-6$^{biotin,His}$ and L-RA07 was observed (Figure 14C). These results demonstrated that D-IL-6$^{biotin,His}$ has mirror-image biological functions of L-IL-6, and can be used for mirror-image screening.

**VHH Library Screening Using D-IL-6 as the Target Protein.** With functional D-IL-6$^{biotin,His}$ in hand, preliminary screening of a T7 phage library was performed to explore IL-6-binding agent(s). The VHH of antibodies recognize specific antigens via three unique complementarity-determining regions (CDRs).[47] In the VHH library, random sequences were positioned at CDR1 and CDR3.[48] CDR3 contained a 9, 12, 15, or 18-residue random sequence, while CDR1 contained a 2-residue random sequence (Figure 15A). Four rounds of biopanning of these phage libraries on immobilized D-IL-6$^{biotin,His}$ were monitored by ELISA. After the second round of biopanning, binding of the phage pools toward D-IL-6 was observed (Figure 15B). D-IL-6-binding phages were enriched in the subsequent third and fourth rounds, yielding seven individual phages. Sequence analysis revealed that five of the VHH clones contained a 12-residue CDR3, while the other two contained a 9- or 15-residue CDR3 (Figure 15A). The 12-residue CDR3 clones were highly homologous, especially at amino acid positions 102 and 109. All of the VHH clones selectively bound to D-IL-6, but not to L-IL-6 (Figure 15C). Although further affinity maturation of these VHH clones is needed to obtain promising lead sequences, this preliminary investigation demonstrated that the applicability of the synthetic D-IL-6 protein for mirror-image screening using phage display technology to develop low-immunogenic D-VHHs[48] against IL-6.

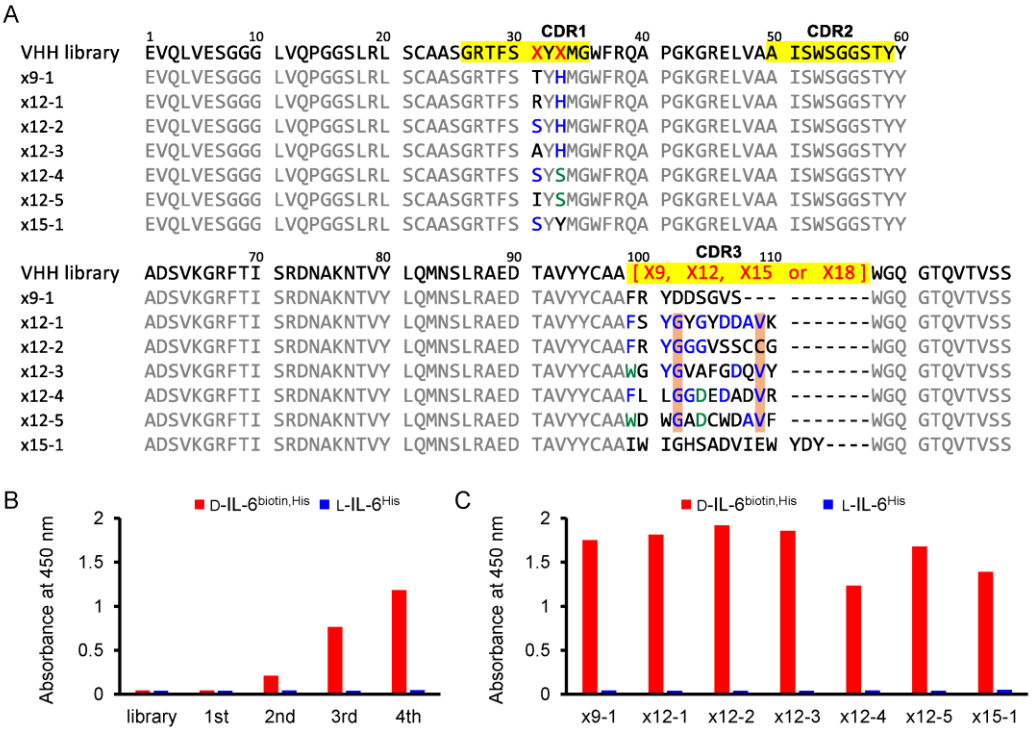


**Figure 15.** Phage display screening of VHH library using D-IL-6$^{biotin,His}$. (A) Sequences of VHH library[48] and identified D-IL-6-binding VHH candidates. Red letters: randomized positions. Blue or green letters: common sequence. (B) ELISA monitoring of the enrichment of D-IL-6-binding phage by four rounds of biopanning. (C) Bioactivities of D-IL-6-binding VHH candidates towards D-IL-6$^{biotin,His}$ or L-IL-6$^{His}$. The binding was determined using ELISA.

In this study, the author established a protocol to synthesize D-IL-6 with applicability for mirror-image screening to explore potential therapeutic agents. The chemical synthesis of IL-6$^{His}$ from six peptide segments was achieved with an N-to-C sequential NCL strategy using multiple Trt-$K_{10}$ solubilizing groups. Folded, synthetic L-IL-6$^{His}$ exhibited structural and functional properties identical to recombinant IL-6. The author also synthesized biotinylated D-IL-6 (D-IL-6$^{biotin,His}$) using the established synthetic protocol of L-IL-6$^{His}$. The symmetrical structural and binding properties of folded D-IL-6$^{biotin,His}$ and L-IL-6 were confirmed by spectroscopic and binding analyses. Preliminary screening of a T7 phage VHH library using D-IL-6$^{biotin,His}$ as a target protein led to the identification of several D-IL-6-binding VHH hit sequences with high homology.

Experimental Section

**General Procedure of Peptide Synthesis.** All reagents and solvents were purchased from Watanabe Chemical Industries, Ltd. (Hiroshima, Japan), Kokusan Chemical Industries, Ltd. (Kanagawa, Japan), Sigma-Aldrich JAPAN (Tokyo, Japan), Wako Pure Chemical Industries, Ltd. (Osaka, Japan), Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan) or Nacalai Tesque, Inc. (Kyoto, Japan). For analytical high-performance liquid chromatography (HPLC), a Cosmosil 5C18-AR300 column (4.6 × 250 mm, Nacalai Tesque, Inc.) was employed with a linear gradient of $CH_3CN$ containing 0.05% (v/v) TFA at a flow rate of 1 mL/min (25 °C). The products were detected by UV absorbance at 220 nm. For preparative HPLC, a Cosmosil 5C18-AR300 column (20 × 250 mm, Nacalai Tesque, Inc.) or a Cosmosil 5C4-AR300 column (20 × 150 mm, Nacalai Tesque, Inc.) was employed with a linear gradient of $CH_3CN$ containing 0.05% TFA at a flow rate of 8 mL/min (room temperature). All peptides were characterized by ESI-MS (LCMS-2020, Shimadzu).

**Solid-Phase Peptide Synthesis.** Standard Fmoc-based solid-phase peptide synthesis (Fmoc-SPPS) was performed using an automatic peptide synthesizer (CYCROSS, DFC, Japan). The following side chain protected amino acids were employed: Arg(Pbf), Asn(Trt), Asp(O*t*-Bu), Cys(Trt), Gln(Trt), Glu(O*t*-Bu), His(Trt), Lys(Boc), Ser(*t*-Bu), Thr(*t*-Bu), and Tyr(*t*-Bu). Fmoc protected amino acids (5 equiv) were coupled using HBTU (5 equiv), HOBt·$H_2O$ (5 equiv) and (*i*-Pr)$_2$NEt (10 equiv) in DMF for 60 min twice. Fmoc protection group was deprotected by 20% piperidine in DMF for 4 min twice unless otherwise stated.

**Microwave-Assisted Solid-Phase Peptide Synthesis.** Microwave-assisted peptide synthesis was performed on a Liberty BLUE (CEM Japan) peptide synthesizer. The following side chain protected amino acids were employed: Arg(Pbf), Asn(Trt), Asp(O*t*-Bu), Cys(Trt), Gln(Trt), Glu(O*t*-Bu), His(Trt), Lys(Boc), Ser(*t*-Bu), Thr(*t*-Bu), Trp(Boc) and Tyr(*t*-Bu). The peptide chain was elongated using the general protocols.[39] Briefly, Fmoc protected amino acids (5 equiv) except for Arg(Pbf) and His(Trt) were coupled using DIC (5 equiv) and Oxyma pure (5 equiv) in DMF at 90 °C for 2 min. Fmoc protection group was deprotected by 20% piperidine in DMF for 1 min at 90 °C.

**[Cys$^{12}$]-L-IL-6$^{12–42}$ (L-3).** C-terminal arginine tag was manually loaded on Rink amide resin (210 mg, 0.10 mmol) by Fmoc-Arg(Pbf)-OH (320 mg, 0.50 mmol), Oxyma Pure (71 mg, 0.50 mmol) and DIC (78 μL, 0.50 mmol) in DMF (1.8 mL) for 2 h. Dbz linker was then loaded on the resin by Fmoc-Dbz-OH (110 mg, 0.30 mmol), HATU (110 mg,

0.30 mmol), and (*i*-Pr)$_2$NEt (110 µL, 0.60 mmol) in DMF (1.8 mL) at 37 °C for 3 h. The first amino acid on the Dbz linker was manually loaded by Fmoc-Thr(*t*-Bu)-OH (200 mg, 0.50 mmol), HATU (190 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 µL, 1.0 mmol) in DMF (1.8 mL) for 1 h. The peptide sequence was constructed by standard Fmoc-SPPS. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The crude products were purified by preparative HPLC to afford the desired peptide L-**3** (83 mg, 16% yield from resin). MS (ESI): calcd for $C_{169}H_{282}N_{56}O_{49}S$: 3914.52; observed: $[M+6H]^{6+}$ *m/z* = 653.40, $[M+5H]^{5+}$ *m/z* = 783.85, $[M+4H]^{4+}$ *m/z* = 979.60, $[M+3H]^{3+}$ *m/z* = 1305.80, $[M+3H+TFA]^{3+}$ *m/z* = 1343.80.

**L-IL-6$^{43–71}$ (L-4).** C-terminal arginine tag was manually loaded on Rink amide resin (210 mg, 0.10 mmol) by Fmoc-Arg(Pbf)-OH (320 mg, 0.50 mmol), Oxyma Pure (71 mg, 0.50 mmol) and DIC (78 µL, 0.50 mmol) in DMF (1.8 mL) for 2 h. Dbz linker was then loaded on the resin by Fmoc-(*o*-Boc)Dbz-OH (140 mg, 0.30 mmol), HATU (110 mg, 0.30 mmol), and (*i*-Pr)$_2$NEt (110 µL, 0.60 mmol) in DMF (1.8 mL) at 37 °C for 3 h. The first amino acid on the Dbz linker was manually loaded by Fmoc-Gly-OH (150 mg, 0.50 mmol), HATU (190 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 µL, 1.0 mmol) in DMF (1.8 mL) for 1 h. The peptide sequence was constructed by standard Fmoc-SPPS. 5% piperazine/0.1 M Oxyma Pure in NMP/EtOH (90:10) cocktail was used for deprotection of Fmoc group instead of 20% piperidine in DMF. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The crude products were purified by preparative HPLC to afford the desired peptide L-**4** (110 mg, 21% yield from resin). MS (ESI): calcd for $C_{158}H_{269}N_{57}O_{52}S_4$: 3927.48; observed: $[M+4H]^{4+}$ *m/z* = 982.75, $[M+3H]^{3+}$ *m/z* = 1310.00.

**[Cys(Acm)$^{72}$]-L-IL-6$^{72–81}$ (L-5).** C-terminal arginine tag was manually loaded on Rink amide resin (210 mg, 0.10 mmol) by Fmoc-Arg(Pbf)-OH (320 mg, 0.50 mmol), Oxyma Pure (71 mg, 0.50 mmol) and DIC (78 µL, 0.50 mmol) in DMF (1.8 mL) for 2 h. Dbz linker was then loaded on the resin by Fmoc-MeDbz-OH (97 mg, 0.25 mmol), HATU (95 mg, 0.25 mmol), and (*i*-Pr)$_2$NEt (87 µL, 0.50 mmol) in DMF (1.8 mL) at 37 °C for 3 h. The first amino acid on the MeDbz linker was manually loaded by Fmoc-Thr(*t*-Bu)-OH (200 mg, 0.50 mmol), HATU (190 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 µL, 1.0 mmol) in DMF (1.8 mL) for 3 h. The peptide sequence was constructed by standard Fmoc-SPPS. For the coupling of an N-terminal cysteine in peptide L-**5**, Boc-Cys(Acm)-OH (150 mg,

0.50 mmol) was employed with Oxyma Pure (71 mg, 0.50 mmol) and DIC (78 μL, 0.50 mmol) in DMF (1.8 mL) for 2 h. Then, the resin was treated with 50 mM 4-nitrophenyl chloroformate in $CH_2Cl_2$ (10 mL) at 37 °C for 2 h followed by 0.5 M (*i*-Pr)$_2$NEt in DMF (10 mL) for 30 min. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The crude products were purified by preparative HPLC to afford the desired peptide L-**5** (120 mg, 45% yield from resin). MS (ESI): calcd for $C_{85}H_{128}N_{32}O_{25}S$: 2030.22; observed: $[M+3H]^{3+}$ *m/z* = 677.65, $[M+2H]^{2+}$ *m/z* = 1015.85.

**L-IL-6$^{82–112}$ (L-6).** C-terminal arginine tag was manually loaded on Rink amide resin (210 mg, 0.10 mmol) by Fmoc-Arg(Pbf)-OH (320 mg, 0.50 mmol), Oxyma Pure (71 mg, 0.50 mmol) and DIC (78 μL, 0.50 mmol) in DMF (1.8 mL) for 2 h. Dbz linker was then loaded on the resin by Fmoc-Dbz-OH (110 mg, 0.30 mmol), HATU (110 mg, 0.30 mmol), and (*i*-Pr)$_2$NEt (110 μL, 0.60 mmol) in DMF (1.8 mL) at 37 °C for 3 h. The first amino acid on the Dbz linker was manually loaded by Fmoc-Arg(Pbf)-OH (320 mg, 0.50 mmol), HATU (190 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 μL, 1.0 mmol) in DMF (1.8 mL) for 2 h. The peptide sequence was constructed by standard Fmoc-SPPS. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The crude products were purified by preparative HPLC to afford the desired peptide L-**6** (110 mg, 20% yield from resin). MS (ESI): calcd for $C_{199}H_{318}N_{60}O_{56}S$: 4479.16; observed: $[M+5H]^{5+}$ *m/z* = 896.75, $[M+4H]^{4+}$ *m/z* = 1120.80, $[M+3H]^{3+}$ *m/z* = 1494.15.

**[Thz$^{113}$]-L-IL-6$^{113–143}$ (L-7).** Thioester L-**7** was synthesized via side-chain-anchoring strategy.[40] Fmoc-Asp-OAllyl (120 g, 0.30 mmol) was manually loaded onto Rink amide (210 mg, 0.10 mmol) by DIC/Oxyma activation. The peptide sequence was constructed by standard Fmoc-SPPS. Boc-Thz-OH was used for the coupling of an N-terminal amino acid in peptide L-**7**. After washing the resin with $CH_2Cl_2$, DMF and dry $CH_2Cl_2$, the resin was treated with a solution of $Pd(PPh_3)_4$ (100 mg, 0.088 mmol) and $PhSiH_3$ (490 μL, 4.0 mmol) in dry $CH_2Cl_2$ (2.0 mL) for 1 h. Afterwards, the resin was washed with $CH_2Cl_2$, DMF and $CH_2Cl_2$. A solution of ethyl 3-mercaptopropionate (300 μL, 21 mmol), HOBt anhydrous (410 mg, 3.0 mmol), (*i*-Pr)$_2$NEt (650 μL, 3.8 mmol) and DIC (470 μL, 3.0 mmol) in $CH_2Cl_2$/DMF (770 μL, 4:1 by volume) was added, and the reaction was continued for 42 h. The reaction was monitored by LC-MS. Final deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5)

for 2 h. After removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**7** (92 mg, 22% yield from resin). MS (ESI): calcd for $C_{157}H_{264}N_{40}O_{48}S_3$: 3576.25; observed: $[M+4H]^{4+}$ *m/z* = 894.95, $[M+3H]^{3+}$ *m/z* = 1193.10, $[M+2H]^{2+}$ *m/z* = 1789.15.

**[Cys$^{144}$]-L-IL-6$^{144–183}$ (L-8).** By the standard procedure of MW-assisted Fmoc-SPPS, the peptide sequence was constructed from Fmoc-Met-Wang resin (350 mg, 0.25 mmol). Global deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford desired peptide L-**8** (77 mg, 5.5% yield from resin). MS (ESI): calcd for $C_{211}H_{347}N_{61}O_{59}S_3$: 4778.65; observed: $[M+4H]^{4+}$ *m/z* = 1195.65, $[M+3H]^{3+}$ *m/z* = 1593.95.

**His$_6$-PEG$_2$-L-IL-6$^{1–11}$-MPAA (L-9a).** C-terminal arginine tag was manually loaded on Rink amide resin (210 mg, 0.10 mmol) by Fmoc-Arg(Pbf)-OH (320 mg, 0.50 mmol), Oxyma Pure (71 mg, 0.50 mmol) and DIC (78 μL, 0.50 mmol) in DMF (1.8 mL) for 2 h. Dbz linker was then loaded on the resin by Fmoc-Dbz-OH (110 mg, 0.30 mmol), HATU (110 mg, 0.30 mmol), and (*i*-Pr)$_2$NEt (110 μL, 0.60 mmol) in DMF (1.8 mL) at 37 °C for 3 h. The first amino acid on the Dbz linker was manually loaded by Fmoc-Ala-OH (170 mg, 0.50 mmol), HATU (190 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 μL, 1.0 mmol) in DMF (1.8 mL) for 1 h. The peptide sequence and PEG linker was constructed by the standard procedure of standard Fmoc-SPPS. Fmoc-NH-PEG$_2$-DGA-OH (Watanabe Chemical) was used to assemble the PEG linker. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The crude products were purified by preparative HPLC to afford the desired Dbz-peptide L-**2a** (160 mg with small amount of impurities, *ca.* 34% yield from resin). For conversion of the C-terminal Dbz into MPAA thioester, an $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 10 eq.) was added to a solution of 4 mM Dbz-peptide L-**2a** in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0) at –20 °C, the reaction was continued for 30 min. An MPAA buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 100 eq.) was then added to the reaction,

and the reaction was continued for 1 min. The reaction was monitored by analytical HPLC. After the reaction was completed, unreacted reagents were separated by preparative HPLC to afford the desired peptide thioester L-**9a** (67 mg, 18% yield from resin). MS (ESI): calcd for $C_{119}H_{176}N_{34}O_{38}S$: 2722.98; observed: $[M+4H]^{4+}$ *m/z* = 681.75, $[M+3H]^{3+}$ *m/z* = 908.50, $[M+2H]^{2+}$ *m/z* = 1362.25.

**[Cys$^{12}$]-His$_6$-PEG$_2$-L-IL-6$^{1–42}$ (L-10a).** Thioester L-**9a** (46 mg, 13 μmol) and cysteine-peptide L-**3** (61 mg, 12 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 6.0 mL) for 1 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**7a** (85 mg, 84% yield). MS (ESI): calcd for $C_{280}H_{450}N_{90}O_{85}S$: 6469.28; observed: $[M+8H]^{8+}$ *m/z* = 809.50, $[M+7H]^{7+}$ *m/z* = 925.15, $[M+6H]^{6+}$ *m/z* = 1079.10, $[M+5H]^{5+}$ *m/z* = 1294.65, $[M+4H]^{4+}$ *m/z* = 1618.25.

**His$_6$-PEG$_2$-L-IL-6$^{1–42}$ (L-11a).** Peptide L-**10a** (78 mg, 9.3 μmol) was reacted in desulfurization buffer (20 mM 2,2'-azobis[2-(2-imidazolin-2-yl)propane]dihydrochloride] (VA-044), 100 mM 2-mercaptethanesulfonate (MESNa), 250 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 6.5; 19 mL) for 2 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**11a** (66 mg, 85% yield). MS (ESI): calcd for $C_{280}H_{450}N_{90}O_{85}$: 6437.22; observed: $[M+8H]^{8+}$ *m/z* = 805.55, $[M+7H]^{7+}$ *m/z* = 920.50, $[M+6H]^{6+}$ *m/z* = 1073.80, $[M+5H]^{5+}$ *m/z* = 1288.45, $[M+4H]^{4+}$ *m/z* = 1610.25.

**His$_6$-PEG$_2$-L-IL-6$^{1–42}$-MPAA (L-12a).** An $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 190 μL) was added to a solution of Dbz-peptide L-**11a** (63 mg, 7.6 μmol) in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 1.9 mL) at –20 °C, the reaction was continued for 30 min. An MPAA buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 1.9 mL) was then added to the reaction, and the reaction was continued for 1 min. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**12a** (57 mg, 93% yield). MS (ESI): calcd for $C_{275}H_{437}N_{83}O_{85}S$: 6298.08; observed: $[M+8H]^{8+}$ *m/z* = 788.10, $[M+7H]^{7+}$ *m/z* = 900.60, $[M+6H]^{6+}$ *m/z* = 1050.60, $[M+5H]^{5+}$ *m/z* = 1260.45, $[M+4H]^{4+}$ *m/z* = 1575.45.

**His$_6$-PEG$_2$-L-IL-6$^{1–71}$ (L-13a).** Thioester L-**12a** (51 mg, 6.4 μmol) and cysteine-peptide L-**4** (34 mg, 6.6 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 3.2 mL) for 3 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative

HPLC to afford the desired peptide L-**13a** (52 mg, 63% yield). MS (ESI): calcd for $C_{425}H_{698}N_{140}O_{135}S_4$: 10057.34; observed: $[M+10H]^{10+}$ *m/z* = 1006.65, $[M+9H]^{9+}$ *m/z* = 1118.45, $[M+8H]^{8+}$ *m/z* = 1257.90, $[M+7H]^{7+}$ *m/z* = 1437.80, $[M+6H]^{6+}$ *m/z* = 1677.30.

**Trt(OH)-K$_{10}$.** Trt(OH)-$K_{10}$ was synthesized by the identical protocol in the previous report.[28] By the standard procedure of standard Fmoc-SPPS, the peptide sequence was constructed from Fmoc-Gly-Wang resin (592 mg, 0.40 mmol). The trityl moiety was manually loaded by 4-(diphenylhydroxymethyl)benzoic acid (243 mg, 0.80 mmol), HATU (304 mg, 0.80 mmol), and (*i*-Pr)$_2$NEt (279 μL, 1.6 mmol) in DMF (7.0 mL) for 2 h. Global deprotection and cleavage from resin was performed by TFA/$H_2O$ (97.5:2.5) for 1 h. After the removal of the resin by filtration, the crude products were purified by preparative HPLC to afford desired Trt(OH)-$K_{10}$ (664 mg, 58% yield from resin). MS (ESI): calcd for $C_{84}H_{142}N_{22}O_{15}$: 1700.20; observed: $[M+3H]^{3+}$ *m/z* = 567.23, $[M+2H]^{2+}$ *m/z* = 850.60.

**[Cys(Acm)$^{72}$/Cys(Trt-K$_{10}$)$^{82}$]-L-IL-6$^{72–112}$ (L-14).** MeNbz-peptide L-**5** (71 mg, 27 μmol) and cysteine-peptide L-**6** (98 mg, 18 μmol) were reacted in 1,2,4-triazole buffer (2.5 M 1,2,4-triazole, 30 mM TCEP, 6 M guanidine·HCl, 100 mM phosphate buffer, pH 7.1; 18 mL) for 14 h at 37 °C. Then, Trt(OH)-$K_{10}$ (520 mg, 180 μmol) followed by TFA (18 mL) were added and the reaction mixture was incubated for 1 h at room temperature. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**14** (52 mg, 30% yield). MS (ESI): calcd for $C_{335}H_{529}N_{95}O_{89}S_2$: 7375.61; observed: $[M+8H+TFA]^{8+}$ *m/z* = 937.15, $[M+7H+TFA]^{7+}$ *m/z* = 1070.85, $[M+6H]^{6+}$ *m/z* = 1230.15, $[M+5H]^{5+}$ *m/z* = 1476.20, $[M+4H]^{4+}$ *m/z* = 1845.00.

**[Cys(Acm)$^{72}$/Cys(Trt-K$_{10}$)$^{82}$]-L-IL-6$^{72–112}$-MPAA (L-15).** An $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 5.8 μL) was added to a solution of Dbz-peptide L-**14** (2.2 mg, 0.23 μmol) in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 58 μL) at –20 °C, the reaction was continued for 30 min. An MPAA buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 58 μL) was then added to the reaction, and the reaction was continued for 1 min. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**15** (0.75 mg, 40% yield). MS (ESI): calcd for $C_{312}H_{480}N_{76}O_{86}S_3$: 6767.90; observed: $[M+6H]^{6+}$ *m/z* = 1128.90, $[M+5H]^{5+}$ *m/z* = 1354.50, $[M+4H]^{4+}$ *m/z* = 1693.00.

**[Thz$^{113}$]-L-IL-6$^{113–143}$-MPAA (L-16).** Alkyl thioester L-**7** (40 mg, 9.7 μmol) was dissolved in MPAA buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM

phosphate buffer, pH 6.5; 15 mL) for 1 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**16** (17 mg, 43% yield). MS (ESI): calcd for $C_{160}H_{262}N_{40}O_{48}S_3$: 3610.27; observed: $[M+4H]^{4+}$ *m/z* = 903.45, $[M+3H]^{3+}$ *m/z* = 1204.40, $[M+2H]^{2+}$ *m/z* = 1806.10.

**[Cys$^{113}$/Cys$^{144}$]-L-IL-6$^{113–183}$ (L-18).** MPAA thioester **L-16** (19 mg, 4.4 μmol) and cysteine-peptide L-**8** (19 mg, 3.4 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 5.1 mL) for 4 h at 37 °C. Then, the same amount of a deprotection buffer (1 M methoxyamine hydrochloride, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 4.0; 5.1 mL) was added and the reaction mixture was incubated for 8 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**18** (14 mg, 42% yield). MS (ESI): calcd for $C_{362}H_{601}N_{101}O_{105}S_5$: 8208.69; observed: $[M+9H]^{9+}$ *m/z* = 913.00, $[M+8H]^{8+}$ *m/z* = 1027.00, $[M+7H]^{7+}$ *m/z* = 1173.55, $[M+6H]^{6+}$ *m/z* = 1369.30, $[M+5H]^{5+}$ *m/z* = 1642.75.

**[Cys(Acm)$^{72}$/Cys(Trt-K$_{10}$)$^{82}$/Cys$^{113}$/Cys$^{144}$]-L-IL-6$^{72–183}$ (L-19, not isolated).** MPAA thioester **L-15** (0.75 mg, 0.10 μmol) and cysteine-peptide L-**18** (1.0 mg, 0.10 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 100 μL) for 2 h at 37 °C. The reaction was monitored by LC-MS; however, most of the desired peptide L-**19** was not eluted from the analytical HPLC column under the standard conditions.

**L-IL-6$^{72–112}$ (L-20).** C-terminal arginine tag was manually loaded on Rink amide resin (210 mg, 0.10 mmol) by Fmoc-Arg(Pbf)-OH (320 mg, 0.50 mmol), Oxyma Pure (71 mg, 0.50 mmol) and DIC (78 μL, 0.50 mmol) in DMF (1.8 mL) for 2 h. Dbz linker was then loaded on the resin by Fmoc-Dbz-OH (110 mg, 0.30 mmol), HATU (110 mg, 0.30 mmol), and (*i*-Pr)$_2$NEt (110 μL, 0.60 mmol) in DMF (1.8 mL) at 37 °C for 3 h. The first amino acid on the Dbz linker was manually loaded by Fmoc-Arg(Pbf)-OH (320 mg, 0.50 mmol), HATU (190 mg, 0.50 mmol), and (*i*-Pr)$_2$NEt (170 μL, 1.0 mmol) in DMF (1.8 mL) for 1 h. The peptide sequence was constructed by standard Fmoc-SPPS. Final deprotection and cleavage from the resin were performed by TFA/$H_2O$/*m*-cresol/thioanisole/1,2-ethanedithiol (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The crude products were purified by preparative HPLC to afford the desired peptide L-**20** (46 mg, 14% yield from resin). MS (ESI): calcd for $C_{248}H_{384}N_{72}O_{74}S_2$: 5622.35; observed: $[M+6H+TFA]^{6+}$ *m/z* = 956.80, $[M+5H]^{5+}$ *m/z* = 1125.35, $[M+4H]^{4+}$ *m/z* = 1406.40, $[M+3H]^{3+}$ *m/z* = 1875.00.

**His$_6$-PEG$_2$-L-IL-6$^{1–112}$ (L-21a).** An $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 90 μL) was added to a solution of Dbz-peptide L-**13a** (47 mg, 3.6 μmol) in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 900 μL) at –20 °C, the reaction was continued for 30 min. An MPAA buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 900 μL) was then added to the reaction, and the reaction was continued for 1 min. The reaction was monitored by LC-MS. After the reaction was completed, unreacted reagents were separated by preparative HPLC and the desired peptide thioester (with small amount of impurities) was lyophilized. Lyophilized peptide thioester (42 mg) and cysteine-peptide L-**20** (24 mg, 3.6 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 1.8 mL) for 1 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**21a** (42 mg, 65% yield). MS (ESI): calcd for $C_{642}H_{1025}N_{193}O_{204}S_6$: 14903.77; observed: $[M+13H]^{13+}$ *m/z* = 1147.45, $[M+12H]^{12+}$ *m/z* = 1242.65, $[M+11H]^{11+}$ *m/z* = 1355.95, $[M+10H]^{10+}$ *m/z* = 1491.05, $[M+9H]^{9+}$ *m/z* = 1656.95.

**[Cys(Trt-K$_{10}$)$^{43}$/Cys(Trt-K$_{10}$)$^{49}$/Cys(Trt-K$_{10}$)$^{72}$/Cys(Trt-K$_{10}$)$^{82}$]-His$_6$-PEG$_2$-L-IL-6$^{1–112}$ (L-22a).** Peptide L-**21a** (41 mg, 2.3 μmol) and Trt(OH)-$K_{10}$ (650 mg, 230 μmol) in guanidine buffer (6 M guanidine·HCl, 20 mM phosphate buffer, pH 7.5; 2.3 mL) were added to TFA (2.3 mL). After stirring for 1 h, the reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**22a** (47 mg, 71% yield). MS (ESI): calcd for $C_{978}H_{1585}N_{281}O_{260}S_6$: 21632.50; observed: $[M+15H]^{15+}$ *m/z* = 1442.95, $[M+14H]^{14+}$ *m/z* = 1546.15, $[M+14H+TFA]^{14+}$ *m/z* = 1554.40, $[M+13H]^{13+}$ *m/z* = 1665.05, $[M+12H]^{12+}$ *m/z* = 1804.05, $[M+11H]^{11+}$ *m/z* = 1967.95.

**[Cys(Trt-K$_{10}$)$^{43}$/Cys(Trt-K$_{10}$)$^{49}$/Cys(Trt-K$_{10}$)$^{72}$/Cys(Trt-K$_{10}$)$^{82}$/Cys$^{113}$/Cys$^{144}$]-His$_6$-PEG$_2$-L-IL-6$^{1–183}$ (L-23a).** An $NaNO_2$ buffer (400 mM $NaNO_2$, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 41 μL) was added to a solution of Dbz-peptide L-**22a** (47 mg, 1.6 μmol) in activation buffer (6 M guanidine·HCl, 200 mM phosphate buffer, pH 3.0; 410 μL) at –20 °C, the reaction was continued for 30 min. An MPAA buffer (400 mM MPAA, 100 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 410 μL) was then added to the reaction, and the reaction was continued for 1 min. The reaction was monitored by LC-MS. After the reaction was completed, unreacted reagents were separated by preparative HPLC and the desired peptide thioester (with small amount of impurities) was lyophilized. Lyophilized peptide thioester (40 mg) and cysteine-peptide L-**18** (13 mg, 1.4 μmol) were reacted in ligation buffer (100 mM MPAA, 50 mM TCEP,

6 M guanidine·HCl, 200 mM phosphate buffer, pH 7.0; 720 μL) for 2 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**23a** (29 mg, 48% yield). MS (ESI): calcd for $C_{1309}H_{2129}N_{363}O_{360}S_{11}$: 29065.27; observed: $[M+27H]^{27+}$ *m/z* = 1077.60, $[M+26H]^{26+}$ *m/z* = 1118.90, $[M+25H]^{25+}$ *m/z* = 1163.80, $[M+24H]^{24+}$ *m/z* = 1212.20, $[M+23H]^{23+}$ *m/z* = 1265.00.

**[Cys(Trt-K$_{10}$)$^{43}$/Cys(Trt-K$_{10}$)$^{49}$/Cys(Trt-K$_{10}$)$^{72}$/Cys(Trt-K$_{10}$)$^{82}$]-His$_6$-PEG$_2$-L-IL-6$^{1–183}$ (L-24a).** Peptide L-**23a** (19 mg, 0.50 μmol) was reacted in desulfurization buffer (20 mM VA-044, 100 mM MESNa, 250 mM TCEP, 6 M guanidine·HCl, 200 mM phosphate buffer, pH 6.5; 1.0 mL) for 2 h at 37 °C. The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**24a** (10 mg, 54% yield). MS (ESI): calcd for $C_{1309}H_{2129}N_{363}O_{360}S_9$: 29001.15; observed: $[M+27H]^{27+}$ *m/z* = 1075.20, $[M+26H]^{26+}$ *m/z* = 1116.80, $[M+25H]^{25+}$ *m/z* = 1161.45, $[M+24H]^{24+}$ *m/z* = 1209.60, $[M+23H]^{23+}$ *m/z* = 1262.05.

**His$_6$-PEG$_2$-L-IL-6$^{1–183}$ (SH form, L-25a).** Peptide L-**24a** (8.5 mg, 0.23 μmol) was treated with TFA/TIS (460 μL, 95:5 by volume). After 30 min, the solution was concentrated. The residue was dissolved in guanidine buffer (6 M guanidine·HCl, 20 mM phosphate buffer, pH 7.5; 460 μL). The reaction was monitored by LC-MS. The crude products were purified by preparative HPLC to afford the desired peptide L-**25a** (2.0 mg, 34% yield). MS (ESI): calcd for $C_{973}H_{1569}N_{275}O_{304}S_9$: 22272.42; observed: $[M+22H]^{22+}$ *m/z* = 1013.30, $[M+21H]^{21+}$ *m/z* = 1061.50, $[M+20H]^{20+}$ *m/z* = 1114.75, $[M+19H]^{19+}$ *m/z* = 1173.30, $[M+18H]^{18+}$ *m/z* = 1238.45.

**His$_6$-PEG$_2$-L-IL-6$^{1–183}$ (SS form, L-1a).** Lyophilized peptide L-**25a** (1.1 mg, 43 nmol) was dissolved in 10 mM Tris buffer (pH 8.5) containing 6 M guanidine·HCl at the concentration of 10 mg/mL. The solution was diluted 10-fold in 10 mM Tris buffer (pH 8.5) and placed for 21 h at room temperature. The reaction was monitored by analytical HPLC. The reaction mixture was purified by size exclusion chromatography. The purified protein was concentrated using an MWCO 10,000 centrifugal filtration membrane (Millipore, Amicon-Ultra 10 kDa) to afford the desired folded protein L-**1a** (11 nmol, 26% yield). The folding yield was estimated by the concentration of the folded protein calculated by absorbance at 280 nm.[49] MS (ESI): calcd for $C_{973}H_{1565}N_{275}O_{304}S_9$: 22268.38; observed: $[M+22H]^{22+}$ *m/z* = 1013.35, $[M+21H]^{21+}$ *m/z* = 1061.20, $[M+20H]^{20+}$ *m/z* = 1114.30, $[M+19H]^{19+}$ *m/z* = 1172.90, $[M+18H]^{18+}$ *m/z* = 1238.05.

**[D-Cys$^{12}$]-D-IL-6$^{12–42}$ (D-3).** By the identical procedure described for the synthesis of L-

**3**, peptide D-**3** was synthesized (110 mg, 11% yield) from Rink amide resin (430 mg, 0.20 mmol). MS (ESI): calcd for $C_{169}H_{282}N_{56}O_{49}S$: 3914.52; observed: $[M+6H]^{6+}$ *m/z* = 653.30, $[M+5H]^{5+}$ *m/z* = 783.80, $[M+4H]^{4+}$ *m/z* = 979.50, $[M+3H]^{3+}$ *m/z* = 1305.65, $[M+3H+TFA]^{3+}$ *m/z* = 1343.90.

**D-IL-6$^{43–71}$ (D-4).** By the identical procedure described for the synthesis of L-**4**, peptide D-**4** was synthesized (100 mg, 20% yield) from Rink amide resin (210 mg, 0.10 mmol). MS (ESI): calcd for $C_{158}H_{269}N_{57}O_{52}S_4$: 3927.48; observed: $[M+4H]^{4+}$ *m/z* = 982.70, $[M+3H]^{3+}$ *m/z* = 1309.95.

**[D-Thz$^{113}$]-D-IL-6$^{113–143}$ (D-7).** By the identical procedure described for the synthesis of L-**7**, peptide D-**7** was synthesized (120 mg, 20% yield) from Rink amide resin (320 mg, 0.15 mmol). MS (ESI): calcd for $C_{157}H_{264}N_{40}O_{48}S_3$: 3576.25; observed: $[M+4H]^{4+}$ *m/z* = 894.80, $[M+3H]^{3+}$ *m/z* = 1192.80, $[M+2H]^{2+}$ *m/z* = 1788.80.

**[D-Cys$^{144}$]-D-IL-6$^{144–183}$ (D-8).** By the identical procedure described for the synthesis of L-**8**, peptide D-**8** was synthesized (240 mg, 17% yield) from Rink amide resin (440 mg, 0.25 mmol). MS (ESI): calcd for $C_{211}H_{347}N_{61}O_{59}S_3$: 4778.65; observed: $[M+4H]^{4+}$ *m/z* = 1195.55, $[M+3H]^{3+}$ *m/z* = 1593.75.

**biotin-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–11}$-MPAA (D-9b).** By the identical procedure described for the synthesis of L-**9a**, peptide D-**9b** was synthesized (130 mg, 17% yield) from Rink amide resin (430 mg, 0.20 mmol). (+)-Biotin was used for the coupling of N-terminus. MS (ESI): calcd for $C_{129}H_{190}N_{36}O_{40}S_2$: 2949.27; observed: $[M+5H]^{5+}$ *m/z* = 591.05, $[M+4H]^{4+}$ *m/z* = 738.30, $[M+3H]^{3+}$ *m/z* = 983.95, $[M+2H]^{2+}$ *m/z* = 1475.55.

**[D-Cys$^{12}$]-biotin-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–42}$ (D-10b).** By the identical procedure described for the synthesis of L-**10a**, peptide D-**9b** (88 mg, 24 μmol) and D-**3** (110 mg, 21 μmol) were converted into D-**10b** (140 mg, 78% yield). MS (ESI): calcd for $C_{290}H_{464}N_{92}O_{87}S_2$: 6695.58; observed: $[M+8H]^{8+}$ *m/z* = 838.00, $[M+7H]^{7+}$ *m/z* = 957.45, $[M+6H]^{6+}$ *m/z* = 1116.75, $[M+5H]^{5+}$ *m/z* = 1340.00, $[M+4H]^{4+}$ *m/z* = 1674.80.

**biotin-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–42}$ (D-11b).** By the identical procedure described for the synthesis of L-**11a**, peptide D-**10b** (140 mg, 16 μmol) was converted into D-**11b** (130 mg, 91% yield). MS (ESI): calcd for $C_{290}H_{464}N_{92}O_{87}S$: 6663.52; observed: $[M+8H]^{8+}$ *m/z* = 833.70, $[M+7H]^{7+}$ *m/z* = 952.80, $[M+6H]^{6+}$ *m/z* = 1111.45, $[M+5H]^{5+}$ *m/z* = 1333.50, $[M+4H]^{4+}$ *m/z* = 1666.75.

**biotin-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–42}$-MPAA (D-12b).** By the identical procedure described for the synthesis of L-**12a**, peptide D-**11b** (120 mg, 14 μmol) was converted into D-**12b**

(73 mg, 65% yield). MS (ESI): calcd for $C_{285}H_{451}N_{85}O_{87}S_2$: 6524.37; observed: $[M+8H]^{8+}$ *m/z* = 816.35, $[M+7H]^{7+}$ *m/z* = 932.90, $[M+6H]^{6+}$ *m/z* = 1088.20, $[M+5H]^{5+}$ *m/z* = 1305.65, $[M+4H]^{4+}$ *m/z* = 1631.90.

**biotin-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–71}$ (D-13b).** By the identical procedure described for the synthesis of L-**13a**, peptide D-**12b** (72 mg, 8.9 μmol) and D-**3** (50 mg, 9.8 μmol) were converted into D-**13b** (68 mg, 59% yield). MS (ESI): calcd for $C_{435}H_{712}N_{142}O_{137}S_5$: 10283.64; observed: $[M+10H]^{10+}$ *m/z* = 1029.25, $[M+9H]^{9+}$ *m/z* = 1143.40, $[M+8H]^{8+}$ *m/z* = 1286.25, $[M+7H]^{7+}$ *m/z* = 1469.95, $[M+6H]^{6+}$ *m/z* = 1714.75.

**[D-Thz$^{113}$]-D-IL-6$^{113–143}$-MPAA (D-16).** By the identical procedure described for the synthesis of L-**16**, peptide D-**7** (110 mg, 26 μmol) was converted into D-**16** (44 mg, 34% yield). MS (ESI): calcd for $C_{160}H_{262}N_{40}O_{48}S_3$: 3610.27; observed: $[M+4H]^{4+}$ *m/z* = 903.30, $[M+3H]^{3+}$ *m/z* = 1204.25, $[M+2H]^{2+}$ *m/z* = 1805.95.

**[D-Cys$^{113}$/D-Cys$^{144}$]-D-IL-6$^{113–183}$ (D-18).** By the identical procedure described for the synthesis of L-**18**, peptide D-**16** (28 mg, 6.7 μmol) and D-**8** (29 mg, 5.2 μmol) were converted into D-**18** (11 mg, 21% yield). MS (ESI): calcd for $C_{362}H_{601}N_{101}O_{105}S_5$: 8208.69; observed: $[M+9H]^{9+}$ *m/z* = 912.75, $[M+8H]^{8+}$ *m/z* = 1026.85, $[M+7H]^{7+}$ *m/z* = 1173.45, $[M+6H]^{6+}$ *m/z* = 1369.10, $[M+5H]^{5+}$ *m/z* = 1642.60.

**D-IL-6$^{72–112}$ (D-20).** By the identical procedure described for the synthesis of L-**20**, peptide D-**20** was synthesized (61 mg, 9.3% yield) from Rink amide resin (210 mg, 0.10 mmol). MS (ESI): calcd for $C_{248}H_{384}N_{72}O_{74}S_2$: 5622.35; observed: $[M+6H+TFA]^{6+}$ *m/z* = 956.90, $[M+5H]^{5+}$ *m/z* = 1125.15, $[M+4H]^{4+}$ *m/z* = 1406.40, $[M+3H]^{3+}$ *m/z* = 1875.00.

**biotin-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–112}$ (D-21b).** By the identical procedure described for the synthesis of L-**21a**, peptide D-**13b** (64 mg, 5.0 μmol) and D-**20** (31 mg, 4.7 μmol) were converted into D-**21b** (40 mg, 44% yield). MS (ESI): calcd for $C_{652}H_{1039}N_{195}O_{206}S_7$: 15130.06; observed: $[M+12H]^{12+}$ *m/z* = 1261.75, $[M+11H]^{11+}$ *m/z* = 1376.15, $[M+10H]^{10+}$ *m/z* = 1513.80, $[M+9H]^{9+}$ *m/z* = 1682.25, $[M+8H]^{8+}$ *m/z* = 1892.05.

**Trt(OH)-k$_{10}$.** By the identical procedure described for the synthesis of Trt(OH)-$K_{10}$, Trt(OH)-$k_{10}$ was synthesized (930 mg, 41% yield) from Fmoc-Gly-Wang resin (620 mg, 0.80 mmol). MS (ESI): calcd for $C_{84}H_{142}N_{22}O_{15}$: 1700.20; observed: $[M+3H]^{3+}$ *m/z* = 567.17, $[M+2H]^{2+}$ *m/z* = 850.09.

**biotin-[D-Cys(Trt-k$_{10}$)$^{43}$/D-Cys(Trt-k$_{10}$)$^{49}$/D-Cys(Trt-k$_{10}$)$^{72}$/D-Cys(Trt-k$_{10}$)$^{82}$]-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–112}$ (D-22b).** By the identical procedure described for the synthesis of L-**22a**, peptide D-**21b** (38 mg, 2.1 μmol) was converted into D-**22b** (44 mg, 72% yield).

MS (ESI): calcd for $C_{988}H_{1599}N_{283}O_{262}S_7$: 21858.80; observed: $[M+15H]^{15+}$ *m/z* = 1458.10, $[M+14H]^{14+}$ *m/z* = 1562.30, $[M+13H]^{13+}$ *m/z* = 1682.25, $[M+12H]^{12+}$ *m/z* = 1822.40, $[M+11H]^{11+}$ *m/z* = 1988.00.

**biotin-[D-Cys(Trt-k$_{10}$)$^{43}$/D-Cys(Trt-k$_{10}$)$^{49}$/D-Cys(Trt-k$_{10}$)$^{72}$/D-Cys(Trt-k$_{10}$)$^{82}$/D-Cys$^{113}$/D-Cys$^{144}$]-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–183}$ (D-23b).** By the identical procedure described for the synthesis of L-**23a**, peptide D-**22b** (43 mg, 1.5 μmol) and D-**18** (11 mg, 1.1 μmol) were converted into D-**23b** (27 mg, 50% yield). MS (ESI): calcd for $C_{1319}H_{2143}N_{365}O_{362}S_{12}$: 29291.57; observed: $[M+27H]^{27+}$ *m/z* = 1086.20, $[M+26H]^{26+}$ *m/z* = 1127.40, $[M+25H]^{25+}$ *m/z* = 1172.60, $[M+24H]^{24+}$ *m/z* = 1221.80, $[M+23H]^{23+}$ *m/z* = 1274.70.

**biotin-[D-Cys(Trt-k$_{10}$)$^{43}$/D-Cys(Trt-k$_{10}$)$^{49}$/D-Cys(Trt-k$_{10}$)$^{72}$/D-Cys(Trt-k$_{10}$)$^{82}$]-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–183}$ (D-24b).** By the identical procedure described for the synthesis of L-**24a**, peptide D-**23b** (26 mg, 0.69 μmol) was converted into D-**24b** (14 mg, 56% yield). MS (ESI): calcd for $C_{1319}H_{2143}N_{365}O_{362}S_{10}$: 29227.45; observed: $[M+27H]^{27+}$ *m/z* = 1084.15, $[M+26H]^{26+}$ *m/z* = 1125.30, $[M+25H]^{25+}$ *m/z* = 1169.80, $[M+24H]^{24+}$ *m/z* = 1219.10, $[M+23H]^{23+}$ *m/z* = 1271.90.

**biotin-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–183}$ (SH form, D-25b).** By the identical procedure described for the synthesis of L-**25a**, peptide D-**24b** (13 mg, 0.35 μmol) was converted into D-**25b** (3.6 mg, 39% yield). MS (ESI): calcd for $C_{983}H_{1583}N_{277}O_{306}S_{10}$: 22498.71; observed: $[M+22H]^{22+}$ *m/z* = 1024.60, $[M+21H]^{21+}$ *m/z* = 1073.75, $[M+20H]^{20+}$ *m/z* = 1127.15, $[M+19H]^{19+}$ *m/z* = 1186.55, $[M+18H]^{18+}$ *m/z* = 1252.85.

**biotin-(D-His)$_6$-PEG$_2$-D-IL-6$^{1–183}$ (SS form, D-1b).** By the identical procedure described for the synthesis of L-**1a**, peptide D-**25b** (1.2 mg, 47 nmol) was converted into D-**1b** (13 nmol, 29% yield). MS (ESI): calcd for $C_{983}H_{1579}N_{277}O_{306}S_{10}$: 22494.68; observed: $[M+22H]^{22+}$ *m/z* = 1023.45, $[M+21H]^{21+}$ *m/z* = 1072.20, $[M+20H]^{20+}$ *m/z* = 1125.65, $[M+19H]^{19+}$ *m/z* = 1184.80, $[M+18H]^{18+}$ *m/z* = 1250.95.

**FLAG-labeled L-RA07.** By the standard procedure of MW-assisted Fmoc-SPPS, the peptide sequence was constructed from Rink amide resin (53 mg, 0.025 mmol). Global deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford desired FLAG-labeled L-RA07 (1.5 mg, 1.5% yield from resin). MS (ESI): calcd for $C_{155}H_{256}N_{40}O_{58}$: 3607.97;

observed: $[M+3H]^{3+}$ *m/z* = 1203.45, $[M+2H]^{2+}$ *m/z* = 1804.80.

**FLAG-labeled D-RA07.** By the standard procedure of MW-assisted Fmoc-SPPS, the peptide sequence was constructed from Rink amide resin (110 mg, 0.050 mmol). Global deprotection and cleavage from resin was performed by TFA/$H_2O$/*m*-cresol/thioanisole/EDT (80:5:5:5:5) for 2 h. After the removal of the resin by filtration, the crude products were precipitated and washed using ice-cold dry $Et_2O$. The resulting precipitate was dissolved in minimum amount of 50% $CH_3CN$ containing 0.05% TFA. The crude products were purified by preparative HPLC to afford desired FLAG-labeled D-RA07 (5.0 mg, 2.6% yield from resin). MS (ESI): calcd for $C_{155}H_{256}N_{40}O_{58}$: 3607.97; observed: $[M+3H]^{3+}$ *m/z* = 1203.45, $[M+2H]^{2+}$ *m/z* = 1804.80.

**Analysis of Disulfide Formation.** Protein digestion tests were performed to determine the formation of disulfide bonds in synthetic L-IL-$6^{His}$. Either $H_2O$ (1.5 µL, for non-reducing conditions) or 500 mM DTT (1.5 µL, for reducing conditions) was added to 30 µM synthetic L-IL-$6^{His}$ in phosphate-buffered saline (PBS, pH 7.4, 25.5 µL) and then incubated at 95 °C for 5 min. After cooling, 500 mM iodoacetamide solution (3.0 µL) was added to the protein solution and the mixture was incubated at room temperature for 20 min in the dark. Then, Lys-C protease (10 µg/mL, 30 µL; Wako, Osaka, Japan) was added and the mixture was incubated at 25 °C for 12 h. The digested samples were analyzed by LC-MS.

**Measurement of CD Spectra.** Folded L-IL-$6^{His}$ and D-IL-$6^{biotin,His}$ were diluted in PBS (pH 7.4) to adjust the concentration to 8 µM. CD spectra of proteins at 20 °C were recorded using a JASCO J-1500 circular dichroism spectrometer (JASCO, Tokyo, Japan).

**SPR Analysis of L-IL-$6^{His}$ with IL-6R.** SPR analysis of recombinant L-IL-6 and synthetic L-IL-$6^{His}$ binding to human IL-6R was carried out using a Biacore X100 SPR instrument (Cytiva, Tokyo, Japan). Recombinant human IL-6 (cat# cyt-484) was purchased from Prospec-Tany TechnoGene Ltd. (Ness-Ziona, Israel). Recombinant human IL-6 receptor (cat# ILR-H5259) was purchased from ACROBiosystems (Tokyo, Japan). HBS-EP buffer (Cytiva) was used as the running buffer at 25 °C, and 1 M NaCl was used as the regeneration solution. For evaluation of IL-6 proteins, recombinant human IL-6R was immobilized (level 1050 RU) on a CM5 sensor chip using an amine coupling kit containing (1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC)/*N*-hydroxysuccinimide (NHS). Recombinant and synthetic L-IL-6 proteins were evaluated for 3 min of contact time, followed by 3 min of dissociation at a flow rate of 30 µL/min. Binding affinities were determined by fitting a 1:1 binding model from triplicate assays.

**Analysis of Binding Analysis between IL-6 and IL-6-binding Peptide by ELISA.** All wash and dilution processes of ELISAs were performed in PBS (pH 7.4) containing 0.025% Tween 20. Microtiter plates (Greiner, 96-well, high binding, Sigma-Aldrich Japan, Tokyo, Japan) were coated with recombinant L-IL-6, synthetic L-IL-$6^{His}$, or D-IL-$6^{biotin,His}$ in 50 mM sodium carbonate buffer (pH 9.6) (50 µL/well; 30 nM) overnight at 4 °C. After coating, wells were washed three times and blocked for 2 h with 1% Blockace (150 µL/well; MEGMILK SNOW BRAND, Tokyo, Japan). After three washes, either L-RA07 or D-RA07 FLAG-labeled peptide (50 µL/well, 0.01–10 µM) was added and incubated for 1 h at 4 °C. After three washes, a 1:5,000 dilution of anti-FLAG monoclonal antibody (50 µL/well; Sigma-Aldrich Japan) was added and incubated for 1 h at 4 °C. After three washes, a 1:5,000 dilution of anti-mouse IgG (H+L)-horseradish peroxidase (HRP) conjugate (50 µL/well; Promega, Tokyo, Japan) was added and incubated for 1 h at 4 °C. After five washes, 3,3′,5,5′-tetramethylbenzidine (TMB) solution (50 µL/well) was added and incubated for 15 min, then the reaction was stopped using 1 M $H_2SO_4$ (50 µL/well). Absorbance was measured at 450 nm using a microplate absorbance reader (Infinite M Plex, TECAN, Männedorf, Switzerland).

**SPR Analysis of Synthetic IL-6 with IL-6-binding Peptide.** SPR analysis was carried out using a Biacore X100 (Cytiva). HBS-EP buffer (Cytiva) was used as the running buffer at 25 °C, and 1 M NaCl was used as the regeneration solution. To evaluate the binding affinity between IL-6 and RA07 peptide, either synthetic L-IL-$6^{His}$ or D-IL-$6^{biotin,His}$ protein was immobilized on a CM5 sensor chip (L-IL-$6^{His}$: 1291 RU; D-IL-$6^{biotin,His}$: 1049 RU) using an EDC/NHS amine coupling kit. L-RA07 and D-RA07 peptides were evaluated for 3 min of contact time, followed by 6 min of dissociation at a flow rate of 30 µL/min. Binding affinities were determined by fitting a 1:1 binding model from triplicate assays.

**Phage Display Selection.** The author used a T7 phage VHH library, in which a part of CDR1 and CDR3 were randomized using degenerating codons.[48] Nunc Amino Immobilizer Surface plates (Thermo Fischer Scientific) were treated for 1 h with D-IL-$6^{biotin,His}$ in phosphate buffer (pH 8, 100 µL/well) as follows: round 1, 1.0 µg/well; round 2, 0.50 µg/well; and rounds 3 and 4, 0.25 µg/well. Next, the wells were washed three times with PBS containing 0.05% Tween 20 (PBST) and blocked with 10 mM ethanol amine in carbonate buffer (pH 9.6) for 1 h. After three washes with PBST, the VHH libraries were suspended in PBS containing 0.25% bovine serum albumin (BSA) and 0.05% Tween 20, applied to the coated wells (100 µL/well), and incubated for 1 h. After 10 washes with PBST, the bound phages were eluted from the wells with 1% sodium

dodecyl sulfate solution (50 µL/well) for 2 min. The eluted phages were amplified by incubating with the SHuffle5403 strain of *Escherichia coli*[48] in 10 mL of Luria-Bertani medium until bacterial lysis occurred. The amplified phage preparations were then used in the next round of biopanning. After four rounds of biopanning, seven individual clones were randomly selected and sequenced.

**ELISA Phage Binding.** Nunc Amino Immobilizer Surface plates (Thermo Fischer Scientific) were treated for 1 h with D-IL-6$^{biotin,His}$ or L-IL-6$^{His}$ in phosphate buffer (pH 8, 0.10 µg/well in 100 µL). Next, the wells were washed three times with PBST and blocked with 10 mM ethanol amine in carbonate buffer (pH 9.6) for 1 h. After three washes with PBST, the initial phage library, the phage pool after each selection round or individual phages (after the 4th round of biopanning) were suspended in PBS containing 0.25% BSA and 0.05% Tween 20 (100 µL/well), applied to the wells, and incubated for 1 h. After three washes with PBST, a 1:5,000 dilution of anti-T7 Tail Fiber monoclonal antibody (100 µL/well, Merck, Darmstadt, Germany) was added and incubated for 1 h. After five washes with PBST, a 1:5,000 dilution of HRP-conjugated anti-mouse IgG (100 µL/well; Jackson ImmunoResearch, West Grove, PA, USA) was added and incubated for 1 h at 4 °C. After five washes with PBST, TMB solution (100 µL/well) was added and incubated for 10 min, then the reaction was stopped using 1 M $H_2SO_4$ (100 µL/well). Absorbance was measured at 450 nm using a microplate absorbance reader (Infinite M Plex).

References and Footnotes


1. Hirano, T.; Yasukawa, K.; Harada, H.; Taga, T.; Watanabe, Y.; Matsuda, T.; Kashiwamura, S.; Nakajima, K.; Koyama, K.; Iwamatsu, A.; Tsunasawa, S.; Sakiyama, F.; Matsui, H.; Takahara, Y.; Taniguchi, T.; Kishimoto, T. *Nature* **1986**, *324*, 73–76.
2. Cressman, D. E.; Greenbaum, L. E.; DeAngelis, R. A.; Ciliberto, G.; Furth, E. E.; Poli, V.; Taub, R. *Science* **1996**, *274*, 1379–1383.
3. Bettelli, E.; Carrier, Y.; Gao, W.; Korn, T.; Storm, T. B.; Oukka, M.; Weiner, H. L.; Kuchroo, V. K. *Nature* **2006**, *441*, 235–238.
4. Wolsk, E.; Mygind, H.; Grøndahl, T. S.; Pedersen, B. K.; van Hall, G. *Am. J. Physiol. Endocrinol. Metab.* **2010**, *299*, E832–E840.
5. Boulanger, M. J.; Chow, D.-c.; Brevnova, E. E.; Garcia, K. C. *Science* **2003**, *300*, 2101–2104.
6. Schaper, F.; Rose-John, S. *Cytokine Growth Factor Rev.* **2015**, *26*, 475–487.
7. Choy, E. H.; De Benedetti, F.; Takeuchi, T.; Hashizume, M.; John, M. R.; Kishimoto, T. *Nat. Rev. Rheumatol.* **2020**, *16*, 335–345.
8. Scott, L. J. *Drugs* **2017**, *77*, 1865–1879.
9. van Rhee, F.; Fayad, L.; Voorhees, P.; Furman, R.; Lonial, S.; Borghaei, H.; Sokol, L.; Crawford, J.; Cornfeld, M.; Qi, M.; Qin, X.; Herring, J.; Casper, C.; Kurzrock, R. *J. Clin. Oncol.* **2010**, *28*, 3701–3708.
10. Kang, S.; Tanaka, T.; Narazaki, M.; Kishimoto, T. *Immunity* **2019**, *50*, 1007–1023.
11. Silverman, J.; Lu, Q.; Bakker, A.; To, W.; Duguay, A.; Alba, B. M.; Smith, R.; Rivas, A.; Li, P.; Le, H.; Whitehorn, E.; Moore, K. W.; Swimmer, C.; Perlroth, V.; Vogt, M.; Kolkman, J.; Stemmer, W. P. *Nat. Biotechnol.* **2005**, *23*, 1556–1561.
12. Frejd, F.; Gudmundsdotter, L. US Pat., US10633423B2, April 28, 2020.
13. Blesemann, N.; Margerie, D.; Asbrand, C.; Rehberg, M.; Savova, V.; Agueusop, I.; Klemmer, D.; Ding-Pfennigdorff, D.; Schwahn, U.; Dudek, M.; Heyninck, K.; De Tavernier, E.; Cornelis, S.; Kohlmann, M.; Nestle, F. O.; Herrmann, M. *Sci. Transl. Med.* **2023**, *15*, eabq4419.
14. Dingman R.; Balu-Iyer, S. V. *J. Pharm. Sci.* **2019**, *108*, 1637–1654.
15. Kalden, J. R.; Schulze-Koops, H. *Nat. Rev. Rheumatol.* **2017**, *13*, 707–718.
16. Burmester, G. R.; Choy, E.; Kivitz, A.; Ogata, A.; Bao, M.; Nomura, A.; Lacey, S.; Pei, J.; Reiss, W.; Pethoe-Schramm, A.; Mallalieu, N. L.; Wallace, T.; Michalska, M.; Birnboeck, H.; Stubenrauch. K.; Genovese, M. C. *Ann. Rheum. Dis.* **2017**, *76*, 1078–1085.

17. Wells, A. F.; Parrino, J.; Mangan, E. K.; Paccaly, A.; Lin, Y.; Xu, C.; Fan, C.; Graham, N. M. H.; van Hoogstraten, H.; Torri, A. *Rheumatol. Ther.* **2019**, *6*, 339–352.
18. NCT02287922: A phase IIb study for ALX-0061 monotherapy in subjects with rheumatoid arthritis. https://clinicaltrials.gov/study/NCT02287922 (accessed November, 2023).
19. Dintzis, H. M.; Symer, D. E.; Dintzis, R. Z.; Zawadzke, L. E.; Berg, J. M *Proteins: Struct., Funct., Bioinf.* **1993**, *16*, 306–308.
20. Snouwaert, J. N.; Leebeek, F. W. G.; Fowlkes, D. M. *J. Biol. Chem.* **1991**, *266*, 23097–23102.
21. Dela Cruz, C. S.; Lee, Y.; Viswanathan, S. R.; El-Guindy, A. S.; Gerlach, J.; Nikiforow, S.; Shedd, D.; Gradoville, L.; Miler, G. *J. Exp. Med.* **2004**, *199*, 503–514.
22. Reif, A.; Siebenhaar, S.; Troster, A.; Schmalzlein, M.; Lechner, C.; Velisetty, P.; Gottwald, K.; Pohner, C.; Boos, I.; Schubert, V.; Rose-John, S.; Unverzagt, C. *Angew. Chem., Int. Ed.* **2014**, *53*, 12125–12131.
23. Reif, A.; Lam, K.; Weidler, S.; Lott, M.; Boos, I.; Lokau, J.; Bretscher, C.; Monnich, M.; Perkams, L.; Schmalzlein, M.; Graf, C.; Fischer, J.-P.; Lechner, C.; Hallstein, K.; Becker, S.; Weyand, M.; Steegborn, C.; Schultheiss, G.; Rose-John, S.; Garbers, C.; Unverzagt, C. *Angew. Chem., Int. Ed.* **2021**, *60*, 13380–13387.
24. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. *Science* **1994**, *266*, 776–779.
25. Asahina, Y.; Komiya, S.; Ohagi, A.; Fujimoto, R.; Tamagaki, H.; Nakagawa, K.; Sato, T.; Akira, S.; Takao, T.; Ishii, A.; Nakahara, Y.; Hojo, H. *Angew. Chem., Int. Ed.* **2015**, *54*, 8226–8230.
26. Murar, C. E.; Ninomiya, M.; Shimura, S.; Karakus, U.; Boyman, O.; Bode, J. W. *Angew. Chem., Int. Ed.* **2020**, *59*, 8425–8429.
27. Wu, H.; Tan, Y.; Ngai, W. L.; Li, X. *Chem. Sci.* **2023**, *14*, 1582–1589.
28. Tsuda, S.; Mochizuki, M.; Ishiba, H.; Yoshizawa-Kumagaye, K.; Nishio, H.; Oishi, S.; Yoshiya, T. *Angew. Chem., Int. Ed.* **2018**, *57*, 2105–2109.
29. Wan, Q.; Danishefsky, S. J. *Angew. Chem., Int. Ed.* **2007**, *46*, 9248–9252.
30. Tsuda, S.; Masuda, S.; Yoshiya, T. *Org. Biomol. Chem.* **2019**, *17*, 1202–1205.
31. Blanco-Conosa, J. B.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2008**, *47*, 6851–6855.
32. Fang, G.-M.; Li, Y.-M.; Shen, F.; Huang, Y.-C.; Li, J.-B.; Lin, Y.; Cui, H.-K.; Liu, L. *Angew. Chem., Int. Ed.* **2011**, *50*, 7645–7649.
33. Tsuda, S.; Shigenaga, A.; Bando, K.; Otaka, A. *Org. Lett.* **2009**, *11*, 823–826.
34. Agouridas, V.; El Mahdi, O.; Diemer, V.; Cargoët, M.; Monbaliu, J.-C. M.; Melnyk, O. *Chem. Rev.* **2019**, *119*, 7328–7443.

35. Wang, J.-X.; Fang, G.-M.; He, Y.; Qu, D.-L.; Yu, M.; Hong, Z.-Y.; Liu, L. *Angew. Chem., Int. Ed.* **2015**, *54*, 2194–2198.
36. We estimated the solubility of peptide segments based on the theoretical isoelectric point (pI). Four Arg residues were attached for the acidic peptide segments (pI < 7), while one Arg residue was attached for neutral or basic peptide segments (pI ≥ 7).
37. Subiros-Funosas, R.; El-Faham, A.; Albericio, F. *Tetrahedron* **2011**, *67*, 8595–8606.
38. Collins, J. M.; Porter, K. A.; Singh, S. K.; Vanier, G. S. *Org. Lett.* **2014**, *16*, 940–943.
39. Blanco-Canosa, J. B.; Nardone, B.; Albericio, F.; Dawson, P. E. *J. Am. Chem. Soc.* **2015**, *137*, 7197–7209.
40. Ficht, S.; Payne, R. J.; Guy, R. T.; Wong, C.-H. *Chem. - Eur. J.* **2008**, *14*, 3620–3629.
41. For the initial attempt at the preparation of peptide thioester L-**9a**, MeDbz linker was used for the C-terminal activation via a PEG linker. However, the activation of MeDbz-peptide with *p*-nitrophenyl chloroformate unexpectedly failed, probably due to the side reactions around PEG moiety.
42. Bang, D.; Kent, S. B. H. *Angew. Chem., Int. Ed.* **2004**, *43*, 2534–2538.
43. Mamahit, Y. P.; Maki, Y.; Okamoto, R.; Kajihara, Y. *Carbohydr. Res.* **2023**, *531*, 108847.
44. Ejima, D.; Watanabe, M.; Sato, Y.; Date, M.; Yamada, N.; Takahara, Y. *Biotechnol. Bioeng.* **1999**, *62*, 301–310.
45. Tsai, P. L.; Chen, S.-F.; Huang, S. Y. *Rev. Anal. Chem.* **2013**, *32*, 257–268.
46. Kobayashi, T.; Kakui, M.; Shibui, T.; Kitano, Y. *Mol. Biotechnol.* **2011**, *48*, 147–155.
47. Muyldermans, S. *Annu. Rev. Biochem.* **2013**, *82*, 775–797.
48. Aoki, K.; Higashi, K.; Oda, S.; Manabe, A.; Maeda, K.; Morise, J.; Oka, S.; Inuki, S.; Ohno, H.; Oishi, S.; Nonaka, M. Submitted.
49. Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. *Protein Sci.* **1995**, *4*, 2411–2423.

# Conclusions

1. The author established a facile synthetic protocol for mirror-image VHHs (D-VHHs). Both enantiomers of anti-GFP VHH, PMP12A2h1, and $^{111}$In-DTPA-labeled PMP12A2h1 were generated using this synthetic protocol. In the comparative biodistribution study in mice, similar distribution profiles were observed between L-VHH and D-VHH, except for the slightly higher accumulation of D-VHH in the stomach. Immunogenicity studies in mice demonstrated that D-VHH induced significantly less anti-drug antibody (ADA) generation than L-VHH.

2. The author established a screening platform by phage display technology to explore D-VHH-based therapeutics. The author constructed the VHH libraries using a modified PMP12A2h1 framework and degeneration codons for randomization, which were compatible with chemical synthesis. Through the mirror-image screening process using a synthetic mirror-image vascular endothelial growth factor A (D-VEGF-A) protein, the author identified a novel D-VHH, which bound to native L-VEGF-A with sub-micromolar affinity and did not elicit ADA generation in mice.

3. The author investigated the chemical synthesis and *in vitro* assembly of full-length core protein [Cp183(C183A)] of the hepatitis B virus (HBV). Cp183(C183A) was synthesized via convergent and C-to-N direction approaches. The resulting synthetic Cp183(C183A) was successfully refolded to form HBV capsid-like particles.

4. The author investigated the chemical synthesis of interleukin-6 (IL-6). Both enantiomers of IL-6 were synthesized from six peptide segments by an established protocol. The symmetrical binding property of D-IL-6 was demonstrated by functional analysis using IL-6-binding peptides. The resulting D-IL-6 was used for mirror-image screening to identify D-IL-6-binding VHH sequences with high homology.

In summary, the author established a screening process from a D-VHH library using chemically synthetic D-proteins and phage display technology. For the practical application of D-VHHs to protein therapeutics, a scalable synthetic approach as well as optimization and functionalization of the resulting D-VHH candidates need to be investigated. The author also achieved chemical syntheses of several therapeutic targets, such as Cp183(C183A) and IL-6. This work will contribute to discovering novel protein therapeutic candidates from unexplored mirror-image biomolecules.

# Acknowledgements

The author would like to express his sincere and wholehearted appreciation to Professor Hiroaki Ohno (Graduate School of Pharmaceutical Sciences, Kyoto University) and Professor Shinya Oishi (Kyoto Pharmaceutical University) for their kind guidance, helpful advice and encouragement throughout this study.

The author also expresses his heartfelt gratitude to Dr. Shinsuke Inuki, Dr. Norihito Arichi, and Dr. Hiroki Akiba (Graduate School of Pharmaceutical Sciences, Kyoto University) for their constructive discussions and encouragement throughout this study.

The author is profoundly grateful to Dr. Motohiro Nonaka (Graduate School of Medicine, Kyoto University) for his thoughtful guidance, kind mentoring and sincere encouragement.

The author would also like to thank Professor Yohei Katoh (Gemome Editing Innovation Center, Hiroshima University) for their guidance on the preparation of recombinant proteins; Professor Hiroyuki Kimura (Research Center for Experimental Modeling of Human Disease, Kanazawa University) for his assistance with the biodistribution assay; Professor Shogo Oka, Dr. Jyoji Morise, Mr. Katsuaki Higashi, Ms. Sakiho Oda, Ms. Asako Manabe, and Ms. Kayuu Maeda (Graduate School of Medicine, Kyoto University) for their assistance with library construction; Dr. Taku Yoshiya, Dr. Shugo Tsuda (Peptide Institute, Inc.), Ms. Naoko Ogata (Kyoto Pharmaceutical University), and Mr. Jumpei Sasaki (Graduate School of Pharmaceutical Sciences, Kyoto University) for their assistance with the synthesis of hepatitis B virus (HBV) core protein; and Professor Koichi Watashi and Dr. Michiyo Kataoka (National Institute of Infectious Disease) for their analysis of HBV capsid-like particles by transmission electron microscopy.

The supports and advices from Professor Shiroh Futaki and Professor Hideaki Kakeya (Graduate School of Pharmaceutical Sciences, Kyoto University) are greatly appreciated.

The author is also grateful to all colleagues in the Department of Bioorganic Medicinal Chemistry and Department of Biopharmaceutical Chemistry (Graduate School of Pharmaceutical Sciences, Kyoto University), Laboratory of Medicinal Chemistry (Kyoto Pharmaceutical University), and Department of Biological Chemistry (Graduate School of Medicine, Kyoto University) for their valuable comments and assistance.

The author would like to thank the Japan Society for the Promotion of Science (JSPS) and JSPS WISE program “The Graduate Program for Medical Innovation (MIP)” for financial support.

Finally, the author thanks his parents, Tomio and Nobuko Aoki, grandparents, Yuichi and Katsuko Sato, and brother, Yusuke Aoki for their understanding and constant encouragement through this study.

# List of Publications

This study was published in the following papers.

Chapter 1.

Section 1. Mirror-Image Single-Domain Antibody for a Novel Nonimmunogenic Drug Scaffold
Keisuke Aoki, Asako Manabe, Hiroyuki Kimura, Yohei Katoh, Shinsuke Inuki, Hiroaki Ohno, Motohiro Nonaka, and Shinya Oishi
*Bioconjug. Chem.* **2023***, 34,* 2055–2065.

Section 2. Engineering a Low-Immunogenic Mirror-Image VHH against Vascular Endothelial Growth Factor
Keisuke Aoki, Katsuaki Higashi, Sakiho Oda, Asako Manabe, Kayuu Maeda, Jyoji Morise, Shogo Oka, Shinsuke Inuki, Hiroaki Ohno, Shinya Oishi, and Motohiro Nonaka
Submitted.

Chapter 2.

Section 1. Synthesis of the Full-Length Hepatitis B Virus Core Protein and Its Capsid Formation
Keisuke Aoki, Shugo Tsuda, Naoko Ogata, Michiyo Kataoka, Jumpei Sasaki, Shinsuke Inuki, Hiroaki Ohno, Koichi Watashi, Taku Yoshiya, and Shinya Oishi
*Org. Biomol. Chem.* **2024**, DOI: 10.1039/d3ob02099a

Section 2. Chemical Synthesis of Interleukin-6 for Mirror-Image Screening
Keisuke Aoki, Kayuu Maeda, Shinsuke Inuki, Hiroaki Ohno, Motohiro Nonaka, and Shinya Oishi
In preparation.